

RESULT 1	
AA25565	
ID	AA25565 standard; Protein; 379 AA.
XX	
AC	AA25565;
XX	
DT	25-MAR-2003 (updated)
DT	18-JAN-1993 (first entry)
XX	
DE	Beta-TG-M1.
XX	
KM	Transforming growth factor beta; induced; CEF-10; v-src; chicken
KW	embryo; fibroblasts; TGF-beta.
XX	
OS	Mus musculus.
XX	
PN	EP495674-A2.
XX	
PD	22-JUL-1992.
XX	
PF	17-JAN-1992; 92EP-0300429.
XX	
PR	18-JAN-1991; 91US-0642991.
PR	10-JAN-1992; 92US-0816270.
XX	
PA	(BRIM) BRISTOL-MYERS SQUIBB CO.
XX	
PI	Brunner AM, Chinn J, Neubauer MG, Purchio AF;
DR	WPI; 1992-243508/30.
DR	N-PDB; AAQ26421.

PT TGF-beta induced gene family - encodes proteins involved in
 PT growth and differentiation effects of TGF-beta-1
 XX
 XX
 PS Claim 2; Fig 1; 35pp; English.

CC The protein sequence was deduced from the DNA sequence obt. by
 CC screening a cDNA library made from AKR-2B mouse cells induced with
 CC TGF-beta1 and cyclohexamide with two probes from untreated AKR-2B
 CC mRNA and AKR-2B mRNA from cells treated with cyclohexamide and TGF-
 CC beta1. The proteins encoded by hybridising colonies (beta-IG-M1 and
 CC beta-IG-M2) contain 38 Cys residues and are induced by TGF-beta1.
 CC Beta-IG-M1 displays 80 percent homology to the CCF-10 protein
 CC induced by v-src in chicken embryo fibroblasts and is identical
 CC to the protein encoded by cyr61, an immediate early response gene
 CC induced in quiescent BALB 3T3 cells by serum treatment. Residues
 CC 49-56 of beta-IG-M1 conform to the GCGCCXXC motif reported in the
 CC amino half of insulin-like growth factor (IGF) binding proteins.
 CC The C-terminal Cys rich region of beta-IG-M1, -M2 and CCF-10 contain
 CC an amino acid sequence with strong homology to a motif found near the
 CC C-terminal of the malarial circumsporozoite (CS) protein, which is
 CC highly conserved among all species of malarial parasites sequenced
 CC to date (designated region II). This motif is also found in
 CC other proteins which have cell adhesive properties that mediate
 CC cell-cell and cell-extracellular matrix interactions, such as
 CC properdin, thrombospondin, and TRAP. The proteins encoded by
 CC TGF-beta induced genes are likely to be involved in mediation of
 CC the biological effects of TGF-beta relating to cell growth and
 CC differentiation. See also AAR25566.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 CC
 XX
 SQ Sequence 379 AA:

Query Match 100.0%; Score 2103; DB 13; Length 379;
 Best Local Similarity 100.0%; Pred. No. 3.7e-156;
 Matches 379; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSSTRITLAVAVTLHLTRLALSTCPAACHCPLPAKCAPGVGLVDDCGCCCKVCAKOL 60
 DB 1 MSSSTRITLAVAVTLHLTRLALSTCPAACHCPLPAKCAPGVGLVDDCGCCCKVCAKOL 60
 QY 61 NECCSKTOPCDHRTKGLKCNFGASTALKGICRAOSGRPEYNSRTYONGESPONCKHQ 120
 DB 61 NECCSKTOPCDHRTKGLKCNFGASTALKGICRAOSGRPEYNSRTYONGESPONCKHQ 120
 QY 121 CTCIDGAVGCIPLCPQELSLPNLGCNPRLVYVSGCCBEMWCDEDSIKSLDDODDLG 180
 DB 121 CTCIDGAVGCIPLCPQELSLPNLGCNPRLVYVSGCCBEMWCDEDSIKSLDDODDLG 180
 QY 181 LDASEVELTRNNELIAGKSSSLKRLPVFETEPRLVFNPLHAHGOKCIQVOTSMSCSKS 240
 DB 181 LDASEVELTRNNELIAGKSSSLKRLPVFETEPRLVFNPLHAHGOKCIQVOTSMSCSKS 240
 QY 241 CGGISTRTVNDNPECLVKEIRICFVRPGOPVYSSIKKSKTSKTSKTSKTSKTSKTSK 300
 DB 241 CGGISTRTVNDNPECLVKEIRICFVRPGOPVYSSIKKSKTSKTSKTSKTSKTSKTSK 300
 QY 301 CSSVKKYRPRYKCGSCVDGRCTPLQTRTVYMRFRCEDEGEFSSKNVMMIOGCKNYNCPHP 360
 DB 301 CSSVKKYRPRYKCGSCVDGRCTPLQTRTVYMRFRCEDEGEFSSKNVMMIOGCKNYNCPHP 360
 QY 361 NEASFRLYSLFNDIHKFRD 379
 DB 361 NEASFRLYSLFNDIHKFRD 379

RESULT 2

AAE05920 standard; Protein; 379 AA.

AAE05920;

24-SEP-2001 (first entry)

DE Mouse cysteine-rich protein (Cyr61).

XX Cysteine-rich protein; Cyr61; extracellular matrix signalling molecule;
 KW fibroblast secreted protein; Fisp12; connective tissue growth factor;
 KW CTRG; ECM; cell adhesion; cell migration; fibroblast cell proliferation;
 KW angiogenesis; wound healing; integrin receptor; atherosclerosis; tumour;
 KW heart disease; fibrosis; gene therapy; mouse.

XX Mus musculus.

XX Key Location/Qualifiers

FT Region 164..226 /note="Cysteine free region"

FT Domain 224..240 /note="Domain III"

XX W0200155210-A2.

XX 02-AUG-2001.

XX 31-JAN-2001; 2001WO-US03267.

XX 31-JAN-2000; 2000US-0495448.

XX 15-MAY-2000; 2000US-0204364.

XX 06-OCT-2000; 2000US-0238705.

XX (MUN1-) MUNIN CORP.

XX Lau LF, Yeung C, Greenspan JA;

XX WPI: 2001-465561/50.

XX N-PSDB; AAD11220.

XX Novel human cysteine-rich protein 61 (Cyr61) fragment useful in methods

XX for screening for modulators of cell adhesion, fibroblast cell

XX proliferation, angiogenesis and cell migration

XX Example 1; Fig 1; 186pp; English.

CC The invention relates to extracellular matrix (ECM) signalling
 CC molecules involved in cellular response to growth factors. More
 CC particularly the invention is directed to cysteine-rich protein
 CC (Cyr61), and Cyr61-related proteins such as fibroblast secreted
 CC protein (Fisp12) and connective tissue growth factor (CTGF) and
 CC nucleic acid molecules encoding such proteins. The polypeptides
 CC of the invention are members of cysteine-rich secreted protein
 CC family. Human Cyr61 fragment is useful in methods for screening
 CC modulators of cell adhesion, cell migration, fibroblast cell
 CC proliferation, angiogenesis, wound healing and Cyr61-Integrin
 CC receptor interaction. Modulator of Cyr61-Integrin alphavbeta3
 CC interaction is used for the preparation of a medicament for the
 CC treatment of atherosclerosis, heart disease, tumour metastasis,
 CC fibrosis, tumour growth, disorders associated with inadequate
 CC angiogenesis; aberrant granulation tissue development; aberrant
 CC fibroblast growth and wounds. Polynucleotides of the invention
 CC are useful in gene therapy. The present sequence is mouse Cyr61
 CC protein.

XX Sequence 379 AA:

Query Match 100.0%; Score 2103; DB 22; Length 379;
 Best Local Similarity 100.0%; Pred. No. 3.7e-156;
 Matches 379; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSSTRITLAVAVTLHLTRLALSTCPAACHCPLPAKCAPGVGLVDDCGCCCKVCAKOL 60
 DB 1 MSSSTRITLAVAVTLHLTRLALSTCPAACHCPLPAKCAPGVGLVDDCGCCCKVCAKOL 60

QY 61 NECCSKTOPCDHRTKGLKCNFGASTALKGICRAOSGRPEYNSRTYONGESPONCKHQ 120
 DB 61 NECCSKTOPCDHRTKGLKCNFGASTALKGICRAOSGRPEYNSRTYONGESPONCKHQ 120

QY 121 CTCIDGAVGCIPLCPQELSLPNLGCNPRLVYVSGCCBEMWCDEDSIKSLDDODDLG 180
 DB 121 CTCIDGAVGCIPLCPQELSLPNLGCNPRLVYVSGCCBEMWCDEDSIKSLDDODDLG 180

Db 121 CTCIDGAVGCIPLCQELSLPMLGCPNRLVKSQCCCEWVCDSDSIKSDIDDDLLG 180
 QY 181 LDASEVELTRNNELIAIGKSSLKRLPVFGTEPRVLFNPLHAHGOKCIYOTTSMQCSKS 240
 Db 181 LDASEVELTRNNELIAIGKSSLKRLPVFGTEPRVLFNPLHAHGOKCIYOTTSMQCSKS 240
 QY 241 CGTGISTRTVNDNPECRVLKKEIRICEVRPCGQPVYSSLKKGKSKTKKSPPEVRYTAG 300
 Db 241 CGTGISTRTVNDNPECRVLKKEIRICEVRPCGQPVYSSLKKGKSKTKKSPPEVRYTAG 300
 QY 301 CSSVKKYRPKYCGSCVDGRCCPTLOTRTVKKMFRCEDGEMFSKNMMIQSCKCNYNCPHP 360
 Db 301 CSSVKKYRPKYCGSCVDGRCCPTLOTRTVKKMFRCEDGEMFSKNMMIQSCKCNYNCPHP 360
 QY 361 NEASFRLYSLENDIHKFRD 379
 Db 361 NEASFRLYSLENDIHKFRD 379

RESULT 3

ID ABB09201 standard; Protein; 379 AA.
 AC ABB09201;
 XX 08-JUL-2002 (first entry)
 DE Mouse cyr6 CNN family protein sequence SEQ ID NO:11.

Human; small CCN-like growth factor; SCGF; vulnary; osteopathic;
 gene therapy; muscle wasting disease; osteoporosis; wound healing;
 tissue regeneration; angiogenesis.

OS Mus sp.
 XX US2002049304-A1.
 XX 25-APR-2002.
 PD 14-MAY-2001; 2001US-0853625.
 PF 06-JUN-1995; 95US-0468847.
 PR 01-APR-1998; 98US-0053587.
 XX (HAST/) HASTINGS G A.
 PA (ADAM/) ADAMS M D.
 XX Hastings GA, Adams MD;
 DR WPI: 2002-382150/41.
 XX Novel isolated polynucleotide sequence encoding a human small CCN-like
 growth factor, useful for treating muscle wasting disease, and
 osteoporosis -
 PS Disclosure: Fig 2A-D; 33pp; English.
 CC The present invention describes human small CCN-like growth factor
 (SCGF). SCGF has vulnary and osteopathic activities, and can be used
 in gene therapy. The SCGF polypeptides and polynucleotides can be used
 for treating muscle wasting diseases, and osteoporosis, and to stimulate
 wound healing and tissue regeneration, to promote angiogenesis and to
 stimulate proliferation of vascular smooth muscle and endothelial cell
 production. The present sequence represents a CNN family protein which
 is given in comparison with the human SCGF in the exemplification of the
 present invention.

SO Sequence 379 AA;

Query Match 100.0%; Score 2103; DB 23; Length 379;
 Best Local Similarity 100.0%; Pred. No. 3,7e-156;
 Matches 379; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSNFRILAVAVTLLHLTRLALSTCPACACHCLEAPKCAPGVGLVRDGGCKYCAKOL 60
 Db 1 MSSNFRILAVAVTLLHLTRLALSTCPACACHCLEAPKCAPGVGLVRDGGCKYCAKOL 60
 QY 61 NEDCSKTQPCDHTKLECNFGASSYALGICRAOSEGRPCENSRITYONGESFOPNCKHQ 120
 Db 61 NEDCSKTQPCDHTKLECNFGASSYALGICRAOSEGRPCENSRITYONGESFOPNCKHQ 120
 QY 121 CTCIDGAVGCIPLCQELSLPMLGCPNRLVKSQCCCEWVCDSDSIKSDIDDDLLG 180
 Db 121 CTCIDGAVGCIPLCQELSLPMLGCPNRLVKSQCCCEWVCDSDSIKSDIDDDLLG 180
 QY 181 LDASEVELTRNNELIAIGKSSLKRLPVFGTEPRVLFNPLHAHGOKCIYOTTSMQCSKS 240
 Db 181 LDASEVELTRNNELIAIGKSSLKRLPVFGTEPRVLFNPLHAHGOKCIYOTTSMQCSKS 240
 QY 241 CGTGISTRTVNDNPECRVLKKEIRICEVRPCGQPVYSSLKKGKSKTKKSPPEVRYTAG 300
 Db 241 CGTGISTRTVNDNPECRVLKKEIRICEVRPCGQPVYSSLKKGKSKTKKSPPEVRYTAG 300
 QY 301 CSSVKKYRPKYCGSCVDGRCCPTLOTRTVKKMFRCEDGEMFSKNMMIQSCKCNYNCPHP 360
 Db 301 CSSVKKYRPKYCGSCVDGRCCPTLOTRTVKKMFRCEDGEMFSKNMMIQSCKCNYNCPHP 360
 QY 361 NEASFRLYSLENDIHKFRD 379
 Db 361 NEASFRLYSLENDIHKFRD 379

RESULT 4

ID AAM35957 standard; Protein; 381 AA.

AC AAM35957;
 XX 05-MAR-1998 (first entry)

Human monocyte mature differentiation factor.

Human; monocyte; mature; differentiation factor; MMDf; macrophage;
 cancer; immune activator; tissue culture; infectious disease.

OS Homo sapiens.

XX JP09234079-A.

XX 09-SEP-1997.

PF 04-MAR-1996; 96JP-0075236.

PR 04-MAR-1996; 96JP-0075236.

PA (TOYM) TOYOBO KK.

XX WPI: 1997-497320/46.

DR N-PSDB; AAT97142.

XX A monocyte mature differentiation factor - useful for the long term

XX tissue culture of macrophage(s)

PS Claim 9; Page 12-13; 22pp; Japanese.

CC The present sequence represents a monocyte mature differentiation
 CC factor (MMDf) which maintains the life of macrophages for long periods
 CC in liquid culture. MMDf can be used as an anti-cancer agent, an immune
 CC activator and to treat infectious diseases.

SO Sequence 381 AA;

Query Match 92.2%; Score 1938; DB 18; Length 381;
 Best Local Similarity 91.4%; Pred. No. 2,9e-143;
 Matches 350; Conservative 9; Mismatches 18; Indels 6; Gaps 2;

QY 1 MSSTFRLAAVAVTLHLTRIALSTCPAACHPLEARKCAPGVGLVBDGGCCGVCAKOL 60
 DB 1 MSSTFRLAAVAVTLHLTRIALSTCPAACHPLEARKCAPGVGLVBDGGCCGVCAKOL 60
 QY 61 NEDCSKTOPCDHTGELCNFGASSTALKGICRAOSEGRPCENSRITONGESFOPNCKHQ 120
 DB 61 NEDCSKTOPCDHTGELCNFGASSTALKGICRAOSEGRPCENSRITONGESFOPNCKHQ 120
 QY 121 CTCIDGAVGCIPLPQELSLPNLGCPNRLVKGSGCCCEWVCEDESIKDLDDODDL- 178
 DB 121 CTCIDGAVGCIPLPQELSLPNLGCPNRLVKGSGCCCEWVCEDESIKDLDDODDLG 180
 QY 179 --IGLDASEVELTRNNELIAGKSSSLKRLPVFGTEPRVLEPNLHAHGOKCIVOTTSMSQ 236
 DB 181 KELGFDASEVELTRNNELIAGKSSSLKRLPVFGMEPRILYNPL--OGOKCIVOTTSMSQ 238
 QY 237 CSKCGIGISTRTVNDNPECKLVKETRICEVRPGGQVYSSLKKGKCSKTKSPPEVRF 296
 DB 239 CSKCGIGISTRTVNDNPECKLVKETRICEVRPGGQVYSSLKKGKCSKTKSPPEVRF 298
 QY 297 TYACSSVKKYRPRKYGSCVDGRCTPLQTRTVKMRFCEDGEMFSKNMMIOSCKCNYN 356
 DB 299 TYACLSVKKYRPRKYGSCVDGRCTPLQTRTVKMRFCEDGEMFSKNMMIOSCKCNYN 358
 QY 357 CPHNEASFRLYSLFNDIHKFRD 379
 DB 359 CPHNEASFRLYSLFNDIHKFRD 381

RESULT 5

AAB90773 standard; Protein: 381 AA.

AAB90773;

15-JUN-2001 (first entry)

Human shear stress-response protein SEQ ID NO: 46.

Human shear stress-response protein; vascular disease;

arteriosclerosis.

Homo sapiens.

WO200125427-A1.

12-APR-2001.

02-OCT-2000; 2000WO-JP06840.

01-OCT-1999; 99JP-0280976.

(KYOW) KYOMA HAKKO KOGYO KK.

(NOJIMA) NOJIMA H.

Nojima H, Yoshisue H, Obayashi M, Oka T, Kawabata A, Sakurada K;

Kuga T, Sekine S, Nakamura Y, Sugano S;

WPI: 2001-266308/27.

N-PSDB: AAH02896.

DNA sequences, proteins encoded by them and antibodies against them

useful in diagnosis and treatment of vascular disease caused by

arteriosclerosis -

Claim 60; Page 345-346; 678pp; Japanese.

The present invention provides the protein and coding sequences of a

number of human shear stress response proteins. These are useful in the

diagnosis, treatment and screening of vascular diseases caused by

arteriosclerosis, including heart failure, post-PTCA restenosis and

hypertension.

XX SQ Sequence 381 AA;
 Query Match 92.2%; Score 1938; DB 22; Length 381;
 Best Local Similarity 91.4%; Pred. No. 2.9e-143;
 Matches 350; Conservative 9; Mismatches 18; Indels 6; Gaps 2;
 QY 1 MSSTFRLAAVAVTLHLTRIALSTCPAACHPLEARKCAPGVGLVBDGGCCGVCAKOL 60
 DB 1 MSSTFRLAAVAVTLHLTRIALSTCPAACHPLEARKCAPGVGLVBDGGCCGVCAKOL 60
 QY 61 NEDCSKTOPCDHTGELCNFGASSTALKGICRAOSEGRPCENSRITONGESFOPNCKHQ 120
 DB 61 NEDCSKTOPCDHTGELCNFGASSTALKGICRAOSEGRPCENSRITONGESFOPNCKHQ 120
 QY 121 CTCIDGAVGCIPLPQELSLPNLGCPNRLVKGSGCCCEWVCEDESIKDLDDODDL- 178
 DB 121 CTCIDGAVGCIPLPQELSLPNLGCPNRLVKGSGCCCEWVCEDESIKDLDDODDLG 180
 QY 179 --IGLDASEVELTRNNELIAGKSSSLKRLPVFGTEPRVLEPNLHAHGOKCIVOTTSMSQ 236
 DB 181 KELGFDASEVELTRNNELIAGKSSSLKRLPVFGMEPRILYNPL--OGOKCIVOTTSMSQ 238
 QY 237 CSKCGIGISTRTVNDNPECKLVKETRICEVRPGGQVYSSLKKGKCSKTKSPPEVRF 296
 DB 239 CSKCGIGISTRTVNDNPECKLVKETRICEVRPGGQVYSSLKKGKCSKTKSPPEVRF 298
 QY 297 TYACSSVKKYRPRKYGSCVDGRCTPLQTRTVKMRFCEDGEMFSKNMMIOSCKCNYN 356
 DB 299 TYACLSVKKYRPRKYGSCVDGRCTPLQTRTVKMRFCEDGEMFSKNMMIOSCKCNYN 358
 QY 357 CPHNEASFRLYSLFNDIHKFRD 379
 DB 359 CPHNEASFRLYSLFNDIHKFRD 381

RESULT 6

AAU79761 standard; Protein: 381 AA.

AAU79761;

30-JUL-2002 (first entry)

Human Cyr61 protein.

Human; uterine leiomyoma proliferation; uterine leiomyoma formation;

Cyr61; cytosstatic.

Homo sapiens.

WO200226193-A2.

04-APR-2002.

28-SEP-2001; 2001WO-US30783.

29-SEP-2000; 2000US-236887P.

(AMHP) AMERICAN HOME PROD CORP.

Zhang Z, Sampath D, Zhu Y, Winneker R;

WPI: 2002-383245/41.

N-PSDB: ABK48899.

Preventing uterine leiomyoma formation or inhibiting proliferation of

uterine leiomyoma in subject, comprises modulating or increasing the

level of Cyr61 in leiomyoma tissue

Disclosure; Fig 6; 92pp; English.

The present invention relates to a method of inhibiting proliferation

QY 1 MSSSTFRTTAAVATLTLHLTRLTALSTCPAACHCPLAEPKAPGVGLVBDGGCCCKVCAKOL 60
 DB 1 MSSRIARALAVVTLTLHLTRLTALSTCPAACHCPLAEPKAPGVGLVBDGGCCCKVCAKOL 60
 QY 61 NEDCSKTQPCDHTKGLKLECNFGASSTALKGICRAOSEGRPCVNSRIYONGESFQPNCKHQ 120
 DB 61 NEDCSKTQPCDHTKGLKLECNFGASSTALKGICRAOSEGRPCVNSRIYONGESFQPNCKHQ 120
 QY 121 CTCIDGAVGCIPLCPQELSLPNIGCPNPLVAVTGQCCCEWVDEDSIKDPMEDQGLLG 180
 DB 121 CTCIDGAVGCIPLCPQELSLPNIGCPNPLVAVTGQCCCEWVDEDSIKDPMEDQGLLG 180
 QY 179 --LGLDASEVELTRNNELIAGKSSILKRLPVGTGEPVLFNPLHAHGOKCIQVTSMSQ 236
 DB 181 KELGPDASEVELTRNNELIAGKSSILKRLPVGTGEPVLFNPLHAHGOKCIQVTSMSQ 238
 QY 237 CSKSCGTGISTRTVNDNPECLVKEIRICVRPCGQPVYSSLKKGKCSKTKKSPPEVR 296
 DB 239 CSKTCGTGISTRTVNDNPECLVKEIRICVRPCGQPVYSSLKKGKCSKTKKSPPEVR 298
 QY 297 TYAGCSSVKKYRKRYGSCVDGRCCTPQTRTYKMRPCEDGEMFSKNVMMIOSCKCNYN 356
 DB 299 TYAGCSSVKKYRKRYGSCVDGRCCTPQTRTYKMRPCEDGEMFSKNVMMIOSCKCNYN 358
 QY 357 CPHPNESFRLYSLENDIHKFRD 379
 DB 359 CPHANEAAPFFYRLFNDIHKFRD 381

RESULT 8

ABB05438
 ID ABB05438 standard; Protein; 381 AA.

AC ABB05438;
 XX

DT 15-APR-2002 (first entry)
 XX

XX Human Cyr61 protein SEQ ID NO:2.
 XX

XX Human: Cyr61; breast cancer; sex steroid receptor; cytostatic; promoter;
 KW sex steroid response element; cysteine rich heparin binding protein;
 KW cell proliferation; heparin binding epidermal growth factor;
 KW epidermal growth factor; basic fibroblastic growth factor.

XX Homo sapiens.
 OS

XX WO200198359-A2.
 PN

XX 27-DEC-2001.
 PD

XX 21-JUN-2001; 2001WO-US19823.
 PF

XX 21-JUN-2000; 2000US-213182P.
 PR

XX 16-MAY-2001; 2001US-291510P.
 PR

XX (AMHP) AMERICAN HOME PROD CORP.
 PA

XX Sampath D, Zhang Z, Winneker R;
 PI

XX WPI: 2002-147796/19.
 DR

XX N-PDOB; ABA93127, ABA93130.
 DR

XX Regulation of Cyr61 expression and activity for preventing and
 PT inhibiting breast cancer comprises use of a Cyr61 neutralizing
 PT antibody, an anti-sense oligonucleotide and an antibody which -

XX Claim 6; Fig 1; 86pp; English.
 PS

XX The present invention describes a method for the prevention or inhibition
 CC of breast cancer cell proliferation. The method comprises administration
 CC of a compound that inhibits the interaction of a sex steroid receptor
 CC with a sex steroid response element of the Cyr61 (cysteine rich heparin-

CC binding protein) promoter. Cyr61 has cytostatic activity. An antibody (1)
 CC which neutralises Cyr61 can be used to prevent or inhibit breast cancer
 CC cell proliferation by blocking sex steroid induced and growth factor
 CC induced synthesis of Cyr61 DNA, where the growth factor is epidermal.
 CC heparin binding epidermal, or basic fibroblastic growth factor. (1) can be
 CC used to diagnose or stage breast cancer where the level of Cyr61 in a
 CC positive/suspect breast cancer cell is compared to the level in a
 CC normal cell, an increase in the level of Cyr61 compared to the level in
 CC normal tissue indicates the presence of breast cancer. The level of
 CC Cyr61 being determined by exposing the tissues to (1), and an increase
 CC in the level of bound antibody by the suspect/positive cell as compared
 CC to the normal tissue indicates the presence of breast cancer. The present
 CC sequence represents the human Cyr61 protein, which is used in the
 CC exemplification of the present invention.

QY Sequence 381 AA;

QY Query Match 92.28; Score 1938; DB 23; Length 381;

QY Best Local Similarity 91.44; Pred. No. 2,9e-143; Mismatches 18; Indels 6; Gaps 2;

QY Matches 350; Conservative 9; Mismatches 18; Indels 6; Gaps 2;

QY 1 MSSSTFRTTAAVATLTLHLTRLTALSTCPAACHCPLAEPKAPGVGLVBDGGCCCKVCAKOL 60
 DB 1 MSSRIARALAVVTLTLHLTRLTALSTCPAACHCPLAEPKAPGVGLVBDGGCCCKVCAKOL 60
 QY 61 NEDCSKTQPCDHTKGLKLECNFGASSTALKGICRAOSEGRPCVNSRIYONGESFQPNCKHQ 120
 DB 61 NEDCSKTQPCDHTKGLKLECNFGASSTALKGICRAOSEGRPCVNSRIYONGESFQPNCKHQ 120
 QY 121 CTCIDGAVGCIPLCPQELSLPNIGCPNPLVAVTGQCCCEWVDEDSIKDPMEDQGLLG 180
 DB 121 CTCIDGAVGCIPLCPQELSLPNIGCPNPLVAVTGQCCCEWVDEDSIKDPMEDQGLLG 180
 QY 179 --LGLDASEVELTRNNELIAGKSSILKRLPVGTGEPVLFNPLHAHGOKCIQVTSMSQ 236
 DB 181 KELGPDASEVELTRNNELIAGKSSILKRLPVGTGEPVLFNPLHAHGOKCIQVTSMSQ 238
 QY 237 CSKSCGTGISTRTVNDNPECLVKEIRICVRPCGQPVYSSLKKGKCSKTKKSPPEVR 296
 DB 239 CSKTCGTGISTRTVNDNPECLVKEIRICVRPCGQPVYSSLKKGKCSKTKKSPPEVR 298
 QY 297 TYAGCSSVKKYRKRYGSCVDGRCCTPQTRTYKMRPCEDGEMFSKNVMMIOSCKCNYN 356
 DB 299 TYAGCSSVKKYRKRYGSCVDGRCCTPQTRTYKMRPCEDGEMFSKNVMMIOSCKCNYN 358
 QY 357 CPHPNESFRLYSLENDIHKFRD 379
 DB 359 CPHANEAAPFFYRLFNDIHKFRD 381

RESULT 9

AA843987
 ID AAB43987 standard; Protein; 455 AA.

AC AAB43987;
 XX

DT 08-FEB-2001 (first entry)
 XX

XX Human cancer associated protein sequence SEQ ID NO:1432.
 XX

XX Human: cancer associated gene; cancer antigen; detection; cancer;
 KW diagnosis; cytostatic; proliferative; vulnerability; immunomodulator;
 KW antiadhesive; antilastmatic; antilethemic; antitumor; antiviral;
 KW dermatological; neuroprotective; thrombolytic; coagulant; nocrotic;
 KW vasotropic; antiproliferative; antiangiogenic; gene therapy; inflammation;
 KW immune disorder; hematopoietic cell disorder; autoimmune disorder;
 KW allergic reaction; graft versus host disease; organ rejection;
 KW haemostatic; thrombolytic; cardiovascular disorder; infection;
 KW neurological disease; drug screening.

XX Homo sapiens.
 OS

PN W020055350-A1.
 XX 21-SEP-2000.
 PD 08-MAR-2000; 2000MO-US05882.
 XX 12-MAR-1999; 99US-0124270.
 PR (HUMA-) HUMAN GENOME SCI INC.
 PA Rosen CA, Ruben SM.
 XX WPI; 2000-587533/55.
 DR N-PSDB; AAC78196.
 XX
 PT Novel isolated nucleic acids comprising sequences encoding peptides
 useful for treating or diagnosing e.g. cancer -
 XX
 PS Claim 11; Page 2116-2118; 2352pp; English.
 XX
 CC AAC7607 to AAC78448 encode the human cancer associated proteins given
 CC in ABA43398 to ABA44239. The proteins can have activities based on the
 CC tissues and cells the genes are expressed in. Example of activities
 CC include: cytostatic; proliferative; vulnary; immunomodulator;
 CC antidiabetic; antisthmatic; antirheumatic; antithrombotic;
 CC antiinflammatory; antihypertensive; antiallergic; antibacterial; antiviral;
 CC dermatologic; neuroprotective; cardiac; thrombolytic; coagulant;
 CC neurotropic; vasotropic; antiparietal and antitumorogenic; the
 CC polynucleotides and polypeptides can be used for preventing, treating or
 CC ameliorating medical conditions and diagnosing pathological conditions.
 CC Polynucleotides, polypeptides, antibodies, agonists and antagonists from
 CC the present invention may be used to treat immune disorders by activating
 CC or inhibiting the proliferation, differentiation or mobilisation of
 CC immune cells; to treat disorders of haematopoietic cells, autoimmune
 CC disorders, allergic reactions, graft versus host disease and organ
 CC rejection, modulate haemostatic or thrombolytic activity, modulate
 CC inflammation, cancers, cardiovascular disorders, neurological disease and
 CC bacterial or viral infections. The peptides, nucleotides, antibodies, and
 CC agonists and antagonists may be also be used in drug screens. AAC78449 to
 CC AAC78457 and ABA44240 represent sequences used in the exemplification of
 CC the present invention.
 CC
 XX
 XX Sequence 455 AA;
 SQ
 Query Match 92.2%; Score 1938; DB 21; Length 455;
 Best Local Similarity 91.4%; Pred. No. 3.6e-143;
 Matches 350; Conservative 9; Mismatches 18; Indels 6; Gaps 2;
 QY 1 MSSSTPRTLAVALVTLHLTRALSTCPACPCPLEAPKAPGVILVDDGGCCCKVCAKOL 60
 DB 75 MSSRIARALALVTLHLTRALSTCPACPCPLEAPKAPGVILVDDGGCCCKVCAKOL 134
 QY 61 NECCSKTQPCDHTKGLGECNFGASTALKGICRAQSEGRPCENSRITQNESEFQPNCKHQ 120
 DB 135 NECCSKTQPCDHTKGLGECNFGASTALKGICRAQSEGRPCENSRITQNESEFQPNCKHQ 194
 QY 121 CTCIDGAVGCIPLCPQELSLPNLCGPNRLVYKSGGCCCEWVDDDESKINSLDDODL 178
 DB 195 CTCIDGAVGCIPLCPQELSLPNLCGPNRLVYKSGGCCCEWVDDDESKINSLDDODL 254
 QY 179 -LGLDPASEVELTRNNELLAIGKSSILKRLVFGTEPRVLNHLAAGQKCIQVOTSMWQ 236
 DB 255 KELGFAPASEVELTRNNELLAIGKSSILKRLVFGTEPRVLNHLAAGQKCIQVOTSMWQ 312
 QY 237 GSNSSCGTGTSTRTNPNPEKRLVKEIRICVRCGQPVYSSLLKGGKCSSTKSPKPVRF 296
 DB 313 GSNSSCGTGTSTRTNPNPEKRLVKEIRICVRCGQPVYSSLLKGGKCSSTKSPKPVRF 372
 QY 297 TYAGCSVKKYRPRKYGSGCVDGRCTPLQRTVKNRPRCDGEMFNSKNNWMIQSKCNVN 356
 DB 373 TYAGCSVKKYRPRKYGSGCVDGRCTPLQRTVKNRPRCDGEMFNSKNNWMIQSKCNVN 432
 QY 357 CPHNEASFRLYSLFNDIHKFRD 379

DB 433 CPHNEASFRLYSLFNDIHKFRD 455
 RESULT 10
 ID ABA76937 standard; Protein; 381 AA.
 XX ABA76937;
 AC ABA76937;
 XX
 DT 05-NOV-2002 (first entry)
 DE
 XX
 XX Human protein, comprising CYP61, designated SECL.
 KW Human; SEC; NOV; immunosuppressive; hepatotropic;
 KW antiinflammatory; angiogenic-associated disorder; diagnostic;
 KW gene therapy; developmental disorder; immune disease;
 KW signal transduction pathway disorder; metabolic disorder;
 KW feeding disorder; obesity; wasting disorder; neurodegenerative disorder;
 KW Alzheimer's disease; Parkinson's disease; behavioural disorder; allergy;
 KW asthma; atherosclerosis; cardiomyopathy; angina pectoris;
 KW autoimmune disease; retinal disease; cirrhosis; diabetes;
 KW infectious disease; human immunodeficiency virus; HIV; cancer;
 KW hypertension; hypotension; multiple sclerosis; urinary retention;
 KW osteoporosis; Crohn's disease; ulcer; neurological disorder; anxiety;
 KW haemophilia; cirrhosis; immunogen; vaccine.
 XX
 OS Homo sapiens.
 XX
 PN W020025705-A2.
 PD 18-JUL-2002.
 XX
 PF 11-JAN-2002; 2002MO-US00609.
 XX
 XX 11-JAN-2001; 2001US-261013P.
 PR 11-JAN-2001; 2001US-261014P.
 PR 11-JAN-2001; 2001US-261018P.
 PR 11-JAN-2001; 2001US-261026P.
 PR 11-JAN-2001; 2001US-261029P.
 PR 17-AUG-2001; 2001US-313170P.
 PR 10-SEP-2001; 2001US-318410P.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Mezes PS, Rastelli L, Herrmann JL, MacDougall JR, Zhong H;
 PI Casman SJ, Boldog F, Shinkels RA, Gorman L, Crasta OR, Mysore KK;
 PI Folkerts O, Martin GB, Eisen A, Spaderina SK, Vernet CAM, Bergh C;
 PI Szytek KA, DiIppio VA, Zernusen BD, Peyman JA, Ellerman K;
 PI Stone DJ, Grosse WM, Alsobrook JP, Lepley DM, Rieger DK;
 PI Burgess CE, Edinger S;
 XX
 XX WPI; 2002-590675/63.
 DR N-PSDB; ABS59522.
 XX
 PT Human SECX/NOVX polypeptide useful for diagnosing, preventing or
 PT treating disorders associated with aberrant expression or activity of
 PT SECX/NOVX nucleic acids and proteins e.g., diabetes -
 XX
 PS Example 3; Page 9; 443pp; English.
 XX
 CC The invention discloses the isolated human polypeptides, and
 CC polynucleotides encoding them, that have been designated SECX and NOVX.
 CC The polypeptides can be used for treating, or delaying, the onset of an
 CC angiogenic-associated disorder or treating a pathological state in a
 CC subject, preferably a mammal. They can also be used in determining the
 CC presence of, or predisposition to, a disease associated with altered
 CC levels of the polypeptides and polynucleotides of any one of the 12
 CC sequences (SEC1-12) for raising antibodies, for identifying an agent
 CC that binds to, or that modulates the expression or activity of the
 CC polypeptide, for treating or preventing a NOVX-associated disorder
 CC (NOV1-8) and as a pharmaceutical composition comprising the polypeptide,
 CC polynucleotide or the antibody. The polypeptides and polynucleotides are

CC useful in diagnostic applications where their amounts are assessed, or
CC for the manufacture of a medicament (e.g. gene therapy) for treating or
CC preventing disorders or syndromes such as developmental disorders, immune
CC diseases, signal transduction pathway disorders, metabolic disorders,
CC feeding disorders (including obesity), wasting disorders,
CC neurodegenerative disorders (including Alzheimer's disease and
CC Parkinson's disease), behavioural disorders, allergies, asthma,
CC atherosclerosis, cardiomyopathy, angina pectoris, autoimmune diseases,
CC retinal disease, cirrhosis, diabetes, infectious disease (bacterial,
CC fungal, protozoal and viral e.g. human immunodeficiency virus, HIV),
CC cancer (e.g. prostate cancer), hypertension, hypotension, multiple
CC sclerosis, urinary retention, osteoporosis, Crohn's disease, ulcers,
CC neurological disorders (e.g. anxiety), haemophilia or cirrhosis. They
CC may also be used as immunogens to produce antibodies specific for the
CC invention, and as vaccines. Further, they are useful for screening
CC potential agonist and antagonist compounds. The sequences presented in
CC ABG76937-ABG76956 are the human SECT-12 and NOV1-8 proteins.

XX
XX Sequence 381 AA:
SQ

Query Match 91.8%; Score 1930; DB 23; Length 381;
Best Local Similarity 90.9%; Pred. No. 1,2e-142;
Matches 348; Conservative 10; Mismatches 19; Indels 6; Gaps 2;

QY 1 MSSSTFRLAVAVTLHLRLALSTCPACCHCPLKAPGVLVDRDGGCCCKVCAKOL 60
DB 1 MSSRIRALALVVTLLHLRLALSTCPACCHCPLKAPGVLVDRDGGCCCKVCAKOL 60
QY 61 NEDCSKTQPCDHDKGLECNFGASSTALKGICRAQSGRCEYNSRIYONGESFOPNCKHQ 120
DB 61 NEDCSKTQPCDHDKGLECNFGASSTALKGICRAQSGRCEYNSRIYONGESFOPNCKHQ 120
QY 121 CTCIDGAVGCIPLCPQELSLPNLGCNPNRLVAVSGCCBEMWCDEDSIKSLDDQDL-- 178
DB 121 CTCIDGAVGCIPLCPQELSLPNLGCNPNRLVAVSGCCBEMWCDEDSIKSLDDQDLG 180
QY 179 --LGDASEVELTRNNELTAIGKSSLKRLPVGTGTRVLPFLNLAHGOKCIYOTTSMSQ 236
DB 181 KELGFDASEVELTRNNELTAIGKSSLKRLPVGTGTRVLPFLNLAHGOKCIYOTTSMSQ 238
QY 237 CSKSCGTGISTRTVNDNPECRLVKTRICEVRPCGQPVYSLLKGGKCSKTKSPBPVRF 296
DB 239 CSKTCGTGISTRTVNDNPECRLVKTRICEVRPCGQPVYSLLKGGKCSKTKSPBPVRF 298
QY 297 TYAGCSSVKKYRKRYGSCVDGRCCTPLQTRVYKMFRCEDGEMFSKNMMIOSCKCNYN 356
DB 299 TYAGCLSVKKYRKRYGSCVDGRCCTPLQTRVYKMFRCEDGEMFSKNMMIOSCKCNYN 358
QY 357 CPHPNASFRLLYSLENDIHKFRD 379
DB 359 CPHANEAAPFFTRLFNDIHKFRD 381

RESULT 11
AAM35730 standard; Protein; 381 AA.
XX
XX AAM35730;
XX
DT 27-MAR-1998 (first entry)
XX
XX Human cysteine rich protein 61 (Cyr61).
XX
XX Cysteine rich protein 61; Cyr61; human;
XX extracellular matrix signalling molecule; cell adhesion;
XX cell migration; cell proliferation; angiogenesis; chondrogenesis;
XX oncogenesis; haemostasis; wound healing; organ regeneration.
XX
XX Homo sapiens.
XX
XX W09733995-A2.
XX
XX 18-SEP-1997.

XX
XX 14-MAR-1997; 97WO-US04193.
XX
XX 15-MAR-1996; 96US-0013958.
XX
XX (MUN1-) MUNIN CORP.
XX
XX Iau LF;
XX
XX WPI: 1997-470875/43.
XX
XX N-PSDB; AAT94699.
XX
XX Isolated and purified cysteine rich protein 61, Cyr61 - useful to
XX modulate e.g. haemostasis, induce wound healing, promote organ
XX regeneration etc
XX
XX Claim 2; Page 112-113; 133pp; English.

CC This protein sequence comprises human cysteine rich protein 61
CC (Cyr61), an extracellular matrix signalling molecule. Its amino
CC acid sequence was deduced from a human placental cDNA clone (see
CC AAT94699). Cyr61 polypeptides can be expressed in transformed or
CC transfected host cells. Cyr61 can be used to modulate
CC haemostasis, induce wound healing in a tissue, promote organ
CC regeneration, improve tissue grafting or promote bone or prothesis
CC implantation (claimed). It can also be used to screen for a
CC modulator of angiogenesis, chondrogenesis, oncogenesis, cell
CC adhesion, cell migration, cell proliferation, expand a population
CC of undifferentiated haematopoietic stem cells in culture and to
CC screen for a mitogen (claimed). Ex vivo methods for using
CC mammalian extracellular matrix signalling molecules to prepare
CC blood products are also provided.

XX
XX Sequence 381 AA:
SQ

Query Match 91.7%; Score 1929; DB 18; Length 381;
Best Local Similarity 90.9%; Pred. No. 1,5e-142;
Matches 348; Conservative 10; Mismatches 19; Indels 6; Gaps 2;

QY 1 MSSSTFRLAVAVTLHLRLALSTCPACCHCPLKAPGVLVDRDGGCCCKVCAKOL 60
DB 1 MSSRIRALALVVTLLHLRLALSTCPACCHCPLKAPGVLVDRDGGCCCKVCAKOL 60
QY 61 NEDCSKTQPCDHDKGLECNFGASSTALKGICRAQSGRCEYNSRIYONGESFOPNCKHQ 120
DB 61 NEDCSKTQPCDHDKGLECNFGASSTALKGICRAQSGRCEYNSRIYONGESFOPNCKHQ 120
QY 121 CTCIDGAVGCIPLCPQELSLPNLGCNPNRLVAVSGCCBEMWCDEDSIKSLDDQDL-- 178
DB 121 CTCIDGAVGCIPLCPQELSLPNLGCNPNRLVAVSGCCBEMWCDEDSIKSLDDQDLG 180
QY 179 --LGDASEVELTRNNELTAIGKSSLKRLPVGTGTRVLPFLNLAHGOKCIYOTTSMSQ 236
DB 181 KELGFDASEVELTRNNELTAIGKSSLKRLPVGTGTRVLPFLNLAHGOKCIYOTTSMSQ 238
QY 237 CSKSCGTGISTRTVNDNPECRLVKTRICEVRPCGQPVYSLLKGGKCSKTKSPBPVRF 296
DB 239 CSKTCGTGISTRTVNDNPECRLVKTRICEVRPCGQPVYSLLKGGKCSKTKSPBPVRF 298
QY 297 TYAGCSSVKKYRKRYGSCVDGRCCTPLQTRVYKMFRCEDGEMFSKNMMIOSCKCNYN 356
DB 299 TYAGCLSVKKYRKRYGSCVDGRCCTPLQTRVYKMFRCEDGEMFSKNMMIOSCKCNYN 358
QY 357 CPHPNASFRLLYSLENDIHKFRD 379
DB 359 CPHANEAAPFFTRLFNDIHKFRD 381

RESULT 12
AAE05921 standard; Protein; 381 AA.
XX
XX AAE05921;
XX

XX 24-SEP-2001 (first entry)
 XX Human cysteine-rich protein (Cyr61).
 XX Cysteine-rich protein; Cyr61; extracellular matrix signalling molecule;
 XX fibroblast secreted protein; fisp12; connective tissue growth factor;
 XX CTGF; ECM; cell adhesion; cell migration; fibroblast cell proliferation;
 XX angiogenesis; wound healing; integrin receptor; atherosclerosis; tumour;
 XX heart disease; fibrosis; gene therapy; human.
 XX Homo sapiens.
 FH Key Location/Qualifiers
 FT Region 163..229
 FT /note="Cysteine free region"
 FT Domain 212..281
 FT /note="Domain III"
 FT Domain 282..381
 FT /note="Domain IV"
 XX W0200155210-A2.
 XX 02-AUG-2001.
 XX 31-JAN-2001; 2001WO-US03267.
 XX 31-JAN-2000; 2000US-0495448.
 XX 15-MAY-2000; 2000US-0204364.
 XX 06-OCT-2000; 2000US-0238705.
 XX (MUNI-) MUNIN CORP.
 XX Lau LF, Yeung C, Greenspan JA;
 XX WPI: 2001-465561/50.
 XX DR N-PSDB; AAD11221.
 XX Novel human cysteine-rich protein 61 (Cyr61) fragment useful in methods
 PT for screening for modulators of cell adhesion, fibroblast cell
 PT proliferation, angiogenesis and cell migration
 XX Claim 30; Page 171-172; 186pp; English.
 XX The invention relates to extracellular matrix (ECM) signalling
 CC molecules involved in cellular response to growth factors. More
 CC particularly the invention is directed to cysteine-rich protein
 CC (Cyr61), and Cyr61-related proteins such as fibroblast secreted
 CC protein (fisp12) and connective tissue growth factor (CTGF) and
 CC nucleic acid molecules encoding such proteins. The polypeptides
 CC of the invention are members of cysteine-rich secreted protein
 CC family. Human Cyr61 fragment is useful in methods for screening
 CC modulators of cell adhesion, cell migration, fibroblast cell
 CC proliferation, angiogenesis, wound healing and Cyr61-integrin
 CC receptor interaction. Modulator of Cyr61-integrin alphavbeta3
 CC interaction is used for the preparation of a medicament for the
 CC treatment of atherosclerosis, heart disease, tumour metastasis,
 CC fibrosis, tumour growth, disorders associated with inadequate
 CC angiogenesis, aberrant granulation tissue development, aberrant
 CC fibroblast growth and wounds. Polynucleotides of the invention
 CC are useful in gene therapy. The present sequence is human Cyr61
 CC protein.
 XX Sequence 381 AA;
 SO Query Match 91.7%; Score 1929; DB 22; Length 381;
 Best Local Similarity 90.9%; Pred. No. 1.5e-142;
 Matches 348; Conservative 10; Mismatches 19; Indels 6; Gaps 2;
 QY 1 MSSSTFTLAVAVTLHLTLALSTCPAACHCPLEAPKAPGVGLVADGGCCCKVCAKOL 60
 DB 1 MSSRIARALAVTLHLTLALSTCPAACHCPLEAPKAPGVGLVADGGCCCKVCAKOL 60

QY 61 NEDCSKTQPCDHTKGLRCNFGASSALKGLICRAOSEGRPCPEYNSRIYONGSEFPNCKHQ 120
 DB 61 NEDCSKTQPCDHTKGLRCNFGASSALKGLICRAOSEGRPCPEYNSRIYONGSEFPNCKHQ 120
 QY 121 CTCIDGAVGCTPLCPQELSLPNLCCPNRLVYKSGGCCCEWVCDSDSIKSLDDQDL-- 178
 DB 121 CTCIDGAVGCTPLCPQELSLPNLCCPNRLVYKSGGCCCEWVCDSDSIKSLDDQDL 180
 QY 179 --LGLDSEVHLTRNNELIAGKSSLRPLVFETERRVLEPNLHAIGOKCIYOTYSWQ 236
 DB 181 KELGFDASEVELTRNNELIANGKRSRLRPLVFEMEPRILLYNPL--OGOKCIYOTYSWQ 238
 QY 237 CSKSCGTGISTRYVNDNPECRVLYKETRICVEYRPGGOPYVSSLKGGKSKTKSPPEVRF 296
 DB 239 CSKTCGTGISTRYVNDNPECRVLYKETRICVEYRPGGOPYVSSLKGGKSKTKSPPEVRF 298
 QY 297 TYAGCCSVKKYRPYKCGSCVDGRCTPLQRTYVARRRCEDEGEMFSKNVMIOGCKNYN 356
 DB 299 TYAGCLSVKKYRPYKCGSCVDGRCTPLQRTYVARRRCEDEGEMFSKNVMIOGCKNYN 358
 QY 357 CPHPEASFRLYSLFNDIHKFRD 379
 DB 359 CPHANEAFPEYRLFNDIHKFRD 381
 RESULT 13
 ID ABB09202
 XX ABB09202 standard; Protein: 374 AA.
 XX ABB09202;
 XX 08-JUL-2002 (first entry)
 XX HCGF CNN family protein sequence SEQ ID NO:12.
 XX Human; small CCN-like growth factor; SCGF; vulnary; osteopathic;
 KW gene therapy; muscle wasting disease; osteoporosis; wound healing;
 KW tissue regeneration; angiogenesis.
 XX Unidentified.
 XX US2002049304-A1.
 XX 25-APR-2002.
 XX 14-MAY-2001; 2001US-0853625.
 XX 06-JUN-1995; 95US-0468847.
 XX 01-APR-1998; 98US-0053587.
 XX (HAST/) HASTINGS G. A.
 XX (ADAM/) ADAMS M. D.
 XX Hastings GA, Adams MD;
 DR WPI: 2002-382150/41.
 XX Novel isolated polynucleotide sequence encoding a human small CCN-like
 PT growth factor, useful for treating muscle wasting disease, and
 PT osteoporosis -
 XX Disclosure; Fig 2A-D; 33pp; English.
 XX The present invention describes human small CCN-like growth factor
 CC (SCGF). SCGF has vulnary and osteopathic activities, and can be used
 CC in gene therapy. The SCGF polypeptides and polynucleotides can be used
 CC for treating muscle wasting diseases, and osteoporosis, and to stimulate
 CC wound healing and tissue regeneration, to promote angiogenesis and to
 CC stimulate proliferation of vascular smooth muscle and endothelial cell
 CC production. The present sequence represents a CNN family protein which
 CC is given in comparison with the human SCGF in the exemplification of the
 CC present invention.

SQ Sequence 374 AA;
 Query Match 86.5%; Score 1819.5; DB 23; Length 374;
 Best Local Similarity 88.3%; Pred. No. 5.1e-134;
 Matches 332; Conservative 10; Mismatches 27; Indels 7; Gaps 3;

QY 1 MSSSTFRTLAVALVTLHLTRALSTCPACCHGLEAPKCAPGVGLVRDGGCCCKVCAKOL 60
 1 MSSRIARELALVTLHLTRVGLSTCPADCHGLEAPKCAPGVGLVRDGGCCCKVCAKOL 60
 DB 61 NECCSKTOPCDHRTKGLGECNFGASTALKGICRAQSEGRCEVNSRIYONGESFOPNCKHQ 120
 61 NECCSKTOPCDHRTKGLGECNFGASTALKGICRAQSEGRCEVNSRIYONGESFOPNCKHQ 120
 DB 121 CTCIDGAVG-CIPLCPOEISLPNIGCPNRLVYVSGCCBEWVCDSDSDDDDL- 178
 121 CTCIDGAVG-CIPLCPOEISLPNIGCPNRLVYVSGCCBEWVCDSDSDSDDDDL 180
 DB 179 --LGIDASEVELTRNNELIAIGSSSLKRLPVFGTEPRVLFNPLAHAGOKCIYQTTSMS 235
 181 GKGLGFDASEVELTRNNELIAVAGSSSLKRLPVFGEMPRILYPL--QGOKCIYQTTSMS 238
 QY 236 QCSKSCGTGISTRTVNDNPECRILVKEIRICEVRPCGOPYVSSLKKGKCKSKTKKSPPEVR 295
 239 QCSKSCGTGISTRTVNDNPECRILVKEIRICEVRPCGOPYVSSLKKGKCKSKTKKSPPEVR 298
 DB 296 FTYAGGSSVKKYKPKYKCGSCVDGRCTPLOTRTVYKMRPFCEDEGEMFSKNVMIQSKCN 355
 299 FTYAGGSSVKKYKPKYKCGSCVDGRCTPLOTRTVYKMRPFCEDEGEMFSKNVMIQSKCN 358
 QY 356 NCPHPNEASFRLYSLE 371
 359 NCPHPNEASFRLYSLE 374
 DB

RESULT 14
 AAR90919 standard; Protein: 375 AA.
 XX AAR90919;
 AC AAR90919;
 DT 25-MAR-2003 (updated)
 DT 25-JUN-1996 (first entry)
 XX
 DE Connective tissue growth factor-2.
 XX
 KM CTGF-2; connective tissue growth factor-2; secreted protein;
 KM cartilaginous growth; skeletal; embryo; cell growth; morphogenesis;
 KM insulin-like growth factor; fibroblast growth factor; cry61.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..24
 FT Protein /label= signal_peptide
 FT /label= mature_protein
 XX
 PN WO9601896-A1.
 XX
 PD 25-JAN-1996.
 PF 12-JUL-1994; 94MO-US07736.
 PR 12-JUL-1994; 94MO-US07736.
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Adams MD, Li H;
 XX
 DR WPI: 1996-097626/10.
 DR N-PSDB; AAT12653.
 XX

PT Connective tissue growth factor-2 and DNA encoding it - useful to
 PT enhance the repair of connective and support tissue, and to enhance
 PT wound healing
 XX
 XX Claim 1; Fig 1A-C; 46pp; English.
 PS
 CC Connective tissue growth factor-2 (CTGF-2) is encoded by a cDNA
 CC (AAT12653) isolated from a human foetal lung cDNA library. The CTGF
 CC polypeptides are structurally and functionally related to a family
 CC of growth factors which include IGF (insulin-like growth factor)
 CC PdgF (platelet-derived growth factor), and Pgf (fibroblast growth
 CC factor). CTGF-2 exhibits 89 percent identity and 93 percent similarity
 CC to Cry61. Cry61 is a growth factor-inducible immediate early gene
 CC initially identified in serum-stimulated mouse fibroblasts. It encodes
 CC a member of an emerging family of secreted proteins which are also a
 CC group of cysteine-rich proteins. This group of GFs are important for
 CC normal growth, differentiation, morphogenesis of the cartilaginous
 CC skeleton of an embryo and cell growth.
 CC (updated on 25-MAR-2003 to correct PN field.)
 XX

SQ Sequence 375 AA;
 Query Match 86.5%; Score 1819.5; DB 17; Length 375;
 Best Local Similarity 88.3%; Pred. No. 5.1e-134;
 Matches 332; Conservative 10; Mismatches 27; Indels 7; Gaps 3;

QY 1 MSSSTFRTLAVALVTLHLTRALSTCPACCHGLEAPKCAPGVGLVRDGGCCCKVCAKOL 60
 1 MSSRIARELALVTLHLTRVGLSTCPADCHGLEAPKCAPGVGLVRDGGCCCKVCAKOL 60
 DB 61 NECCSKTOPCDHRTKGLGECNFGASTALKGICRAQSEGRCEVNSRIYONGESFOPNCKHQ 120
 61 NECCSKTOPCDHRTKGLGECNFGASTALKGICRAQSEGRCEVNSRIYONGESFOPNCKHQ 120
 DB 121 CTCIDGAVG-CIPLCPOEISLPNIGCPNRLVYVSGCCBEWVCDSDSDDDDL- 178
 121 CTCIDGAVG-CIPLCPOEISLPNIGCPNRLVYVSGCCBEWVCDSDSDSDDDDL 180
 DB 179 --LGIDASEVELTRNNELIAIGSSSLKRLPVFGTEPRVLFNPLAHAGOKCIYQTTSMS 235
 181 GKGLGFDASEVELTRNNELIAVAGSSSLKRLPVFGEMPRILYPL--QGOKCIYQTTSMS 238
 QY 236 QCSKSCGTGISTRTVNDNPECRILVKEIRICEVRPCGOPYVSSLKKGKCKSKTKKSPPEVR 295
 239 QCSKSCGTGISTRTVNDNPECRILVKEIRICEVRPCGOPYVSSLKKGKCKSKTKKSPPEVR 298
 DB 296 FTYAGGSSVKKYKPKYKCGSCVDGRCTPLOTRTVYKMRPFCEDEGEMFSKNVMIQSKCN 355
 299 FTYAGGSSVKKYKPKYKCGSCVDGRCTPLOTRTVYKMRPFCEDEGEMFSKNVMIQSKCN 358
 QY 356 NCPHPNEASFRLYSLE 371
 359 NCPHPNEASFRLYSLE 374
 DB

RESULT 15
 AAY31620 standard; Protein: 375 AA.
 XX AAY31620;
 AC AAY31620;
 DT 02-NOV-1999 (first entry)
 DT
 DE Human CTGF-2.
 XX
 KM Connective tissue growth factor-2; CTGF-2; wound healing; bone disorder;
 KM skin disorder; acne; burn; UV damage; stabilisation; tissue implant.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..24
 FT /label= signal_peptide
 XX

FT Protein 25..375
FT Misc-difference 268 /note= "Cys encoded by ICF"
XX
XX US5945300-A.
XX
XX 31-AUG-1999.
XX
XX PD
XX PF 02-JUN-1995; 95US-0459101.
XX
XX PR 02-JUN-1995; 95US-0459101.
XX PR 12-JUL-1994; 94WO-US07736.
XX
XX PA (ADAM/) ADAMS M D.
XX (LTHH/) LI H.
XX
XX PI Adams MD, L1 H;
XX
XX WPI; 1999-508171/42.
XX DR N-PSDB; AA211720.
XX
XX PT Polynucleotides encoding growth factor polypeptides useful for
XX enhancing the repair of connective tissue and support tissue
XX
XX PS Claim 1; Fig 1; 20pp; English.
XX
XX This sequence represents human connective tissue growth factor-2
CC (CTGF-2). CTGF-2 cDNA was isolated from a cDNA library derived from
CC human foetal lung. In one instance, the cDNA was cloned into a
CC baculovirus expression vector, having first been amplified and modified
CC via PCR using primers AA211721 and AA211722. In another instance, the
CC cDNA was cloned into a COS cell expression vector, with prior
CC amplification and modification using PCR primers AA211723 and AA211724.
CC CTGF-2 is structurally and functionally related to a family of growth
CC factors which include IGF (insulin-like growth factor), PDGF
CC (platelet-derived growth factor) and FGF (fibroblast growth factor). This
CC emerging family of cysteine-rich secreted proteins are important for
CC normal growth, differentiation, morphogenesis of the cartilaginous
CC skeleton of an embryo and cell growth. Their functions also include wound
CC healing, tissue repair, implant fixation and stimulating increased bone
CC mass. CTGF-2 may be used to enhance the repair of connective tissue and
CC support tissue and can therefore treat skin disorders e.g., acne, aging,
CC UV damage or burns. CTGF-2 can be used to promote the attachment,
CC fixation and stabilization of tissue implants inserted during
CC reconstructive surgery, and can be used to enhance the healing of
CC external wounds. It can be used in the treatment of injured or depleted
CC bone as it promotes the growth of connective tissue, bone and cementum
CC and stimulates protein and collagen synthesis.
XX
XX
SQ Sequence 375 AA;

Query Match 86.5%; Score 1819.5; DB 20; Length 375;
Best Local Similarity 88.3%; Pred. No. 5.1e-134;
Matches 332; Conservative 10; Mismatches 27; Indels 7; Gaps 3;

OY 1 MSSSTFRTIAVAVLLHLTRALSTCPAACHCPLAPKAPGVGLVDRGCCCKVCAKQL 60
DB 1 MSSRIVELALVYLLHTRVGLSTCPADCHCPLAPKAPGVGLVDRGCCCKVCAKQL 60

OY 61 NEDSKTQPCDHTKGLCECNFGASSTALKGICRAOSEGRPCYNSRIYONGESFOPNCKHQ 120
DB 61 NEDCKRTQPCDHTKGLCECNFGASSTALKGICRAOSEGRPCYNSRIYONGESFOPNCKHQ 120

OY 121 CTCIDGAVG-CIPLCPQELSLPNLGCNPRLVKVSGQCCCEBWCDEDSIKDSLDDQDDL- 178
DB 121 CTCIGMRGACIPLCPQELSLPNLGCNPRLVKVSGQCCCEBWCDEDSIKDPMEDQDL 180

OY 179 ---IGLASEVELTRNNELIIGKSSIKRLPVFTGTERPVFNPLHAHGOKCIYOTISMS 235
DB 181 GKGLGFASSEVELTRNNELIIGKSSIKRLPVFGMEPRILYNPL--OGOKCIYOTISMS 238

OY 236 QCSKSCGTGISTRTYNDNPECLVETRICVPRPGQPVYSSLKGGKCSKTKKSPPEVR 295
DB 236 QCSKTCGTGISTRTYNDNPECLVETRICVPRPGQPVYSSLKGGKCSKTKKSPPEVR 298
OY 296 FTYAGCSSLVKYRKRCGSCVDGRCCTPLQTRTKMRRCGDEGEMFSKNNMIOSKCN 355
DB 299 FTYAGCSSLVKYRKRCGSCVDGRCCTPLQTRTKMRRCGDEGEMFSKNNMIOSKCN 358
OY 356 NCPHNEASFRLYSLF 371
DB 359 NCPHNEASFRLYSLF 374

Search completed: August 5, 2003, 14:05:11
Job time : 40.8974 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 5, 2003, 14:03:52 ; Search time 13.9632 Seconds
(without alignments)
1148.438 Million cell updates/sec

Title: US-09-495-448a-2
Perfect score: 2103
Sequence: 1 MSSSTFRLAAVAVTLHLTR.....PNEASFRLYSIFNDHKFRD 379

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCITUS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2103	100.0	379	1	US-08-468-847B-11
2	2103	100.0	379	4	US-09-142-569-2
3	1938	92.2	381	4	US-09-348-815-2
4	1929	91.7	381	4	US-09-142-569-4
5	1819.5	86.5	374	1	US-08-468-847B-12
6	1812.5	86.2	375	2	US-08-459-101A-2
7	1646.5	78.3	375	1	US-08-468-847B-13
8	958	45.6	347	4	US-09-582-337-2
9	957.5	45.5	348	4	US-09-292-036-3
10	953.5	45.3	348	1	US-08-468-847B-15
11	953.5	45.3	348	4	US-09-142-569-6
12	950	45.2	349	1	US-08-167-628-2
13	950	45.2	349	1	US-08-386-680-2
14	950	45.2	349	1	US-08-459-717-2
15	950	45.2	349	1	US-08-712-302-2
16	950	45.2	349	2	US-08-880-031-2
17	950	45.2	349	3	US-09-054-368-2
18	950	45.2	349	3	US-09-097-179-2
19	950	45.2	349	3	US-09-054-274-2
20	950	45.2	349	3	US-09-080-715-2
21	950	45.2	349	3	US-09-056-704-2
22	950	45.2	349	4	US-09-292-036-4
23	950	45.2	349	4	US-09-253-316-26
24	950	45.2	349	4	US-09-142-569-8
25	950	45.2	349	4	US-09-461-688-2
26	950	45.2	349	5	PCT-US96-08140-2
27	944	44.9	347	4	US-09-187-478-2

28	944	44.9	347	4	US-09-292-036-2	Sequence 2, Appl
29	941.5	44.8	348	1	US-08-468-847B-14	Sequence 14, Appl
30	861	40.9	351	1	US-08-468-847B-16	Sequence 16, Appl
31	839.5	39.9	357	1	US-08-468-847B-17	Sequence 17, Appl
32	839.5	39.9	357	4	US-09-253-316-25	Sequence 25, Appl
33	773	36.8	345	4	US-09-182-145-3	Sequence 3, Appl
34	773	36.8	345	4	US-09-182-145-6	Sequence 4, Appl
35	773	36.8	367	4	US-09-182-145-4	Sequence 6, Appl
36	773	36.8	367	4	US-09-182-145-8	Sequence 8, Appl
37	772	36.7	345	4	US-09-182-145-5	Sequence 5, Appl
38	772	36.7	345	4	US-09-182-145-21	Sequence 21, Appl
39	772	36.7	367	4	US-09-182-145-7	Sequence 7, Appl
40	772	36.7	367	4	US-09-182-145-22	Sequence 22, Appl
41	758.5	36.1	345	4	US-09-182-145-11	Sequence 11, Appl
42	758.5	36.1	367	4	US-09-182-145-12	Sequence 12, Appl
43	622	29.6	339	4	US-09-182-145-36	Sequence 36, Appl
44	622	29.6	354	4	US-09-182-145-37	Sequence 37, Appl
45	622	29.6	354	4	US-09-253-316-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-08-468-847B-11
; Sequence 11, Application US/08468847B
; Patent No. 5780263
; GENERAL INFORMATION:
; APPLICANT: Hastings, Gregg A. and Adams, Mark D.
; TITLE OF INVENTION: Human CCN-Like Growth Factor
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
; ADDRESSEE: CECCHI, STEWART & OLSTEIN
; STREET: 6 BECKER FARM ROAD
; CITY: ROSELAND
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468,847B
; FILING DATE: 6 June 1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: MULLINS, J. G.
; REGISTRATION NUMBER: 33,073
; REFERENCE/DOCKET NUMBER: 325800-442
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 379 AMINO ACIDS
; TYPE: AMINO ACID
; STRANDEDNESS:
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PROTEIN
; US-08-468-847B-11
; Query Match 100.0%; Score 2103; DB 1; Length 379;
; Best Local Similarity 100.0%; Pred. No. 4.2e-170;
; Matches 379; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 MSSSTFRLAAVAVTLHLTRALSTCPACACPLEAPKCAPGVGLVRDCCCKKVCARQL 60
|||||

Db 1 MSSFTPTLAVATLTLTLALSTCPAACHCPLEAPKAPGVGLVRDGGCCCKVCAKOL 60
QY 61 NEDSKTOPCDHTKGLGECNFGASTALKGICRAOSEGRPCENSRITYONGESFOPNCKHQ 120
Db 61 NEDSKTOPCDHTKGLGECNFGASTALKGICRAOSEGRPCENSRITYONGESFOPNCKHQ 120
QY 121 CTCIDGAVGCIPLCPQELSLPNIIGCPNRLVKVSGGCCCEWVDESDIKSLDDODDLG 180
Db 121 CTCIDGAVGCIPLCPQELSLPNIIGCPNRLVKVSGGCCCEWVDESDIKSLDDODDLG 180
QY 181 LDASEVELTRNNELIAGKSSSLKRLPVFETEPRLVFNPLHAHQKCIYOTTSMSQCSKS 240
Db 181 LDASEVELTRNNELIAGKSSSLKRLPVFETEPRLVFNPLHAHQKCIYOTTSMSQCSKS 240
QY 241 CGTISTRTVNDNPECLVETRICVPRCPGQPVYSSLLKGGKCSKTKKSPPEVRYTAG 300
Db 241 CGTISTRTVNDNPECLVETRICVPRCPGQPVYSSLLKGGKCSKTKKSPPEVRYTAG 300
QY 301 CSSVKYRPRYKCGSCVDGRCTPLQRTVYKMRFRCEDEGEMFSKNVMMIOCKCNVNCPPH 360
Db 301 CSSVKYRPRYKCGSCVDGRCTPLQRTVYKMRFRCEDEGEMFSKNVMMIOCKCNVNCPPH 360
QY 361 NEASFRLYSLFNDIHKFRD 379
Db 361 NEASFRLYSLFNDIHKFRD 379

RESULT 2
US-09-142-569-2
; Sequence 2, Application US/09142569
; Patent No. 6413735

GENERAL INFORMATION:

APPLICANT: Lau, Lester F.
TITLE OF INVENTION: Extracellular Matrix Signalling Molecules
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/142,569
FILING DATE: 02-Apr-1999
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Clough, David W.
REGISTRATION NUMBER: 36,107
REFERENCE/DOCKET NUMBER: 28758/33766
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 379 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: misc.feature
OTHER INFORMATION: "Mouse Cyt61 amino acid sequence"

US-09-142-569-2
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
Query Match 100.0%; Score 2103; DB 4; Length 379;
Best Local Similarity 100.0%; Pred. No. 4,2e-170;

Matches 379; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSSFTPTLAVATLTLTLALSTCPAACHCPLEAPKAPGVGLVRDGGCCCKVCAKOL 60
Db 1 MSSFTPTLAVATLTLTLALSTCPAACHCPLEAPKAPGVGLVRDGGCCCKVCAKOL 60
QY 61 NEDSKTOPCDHTKGLGECNFGASTALKGICRAOSEGRPCENSRITYONGESFOPNCKHQ 120
Db 61 NEDSKTOPCDHTKGLGECNFGASTALKGICRAOSEGRPCENSRITYONGESFOPNCKHQ 120
QY 121 CTCIDGAVGCIPLCPQELSLPNIIGCPNRLVKVSGGCCCEWVDESDIKSLDDODDLG 180
Db 121 CTCIDGAVGCIPLCPQELSLPNIIGCPNRLVKVSGGCCCEWVDESDIKSLDDODDLG 180
QY 181 LDASEVELTRNNELIAGKSSSLKRLPVFETEPRLVFNPLHAHQKCIYOTTSMSQCSKS 240
Db 181 LDASEVELTRNNELIAGKSSSLKRLPVFETEPRLVFNPLHAHQKCIYOTTSMSQCSKS 240
QY 241 CGTISTRTVNDNPECLVETRICVPRCPGQPVYSSLLKGGKCSKTKKSPPEVRYTAG 300
Db 241 CGTISTRTVNDNPECLVETRICVPRCPGQPVYSSLLKGGKCSKTKKSPPEVRYTAG 300
QY 301 CSSVKYRPRYKCGSCVDGRCTPLQRTVYKMRFRCEDEGEMFSKNVMMIOCKCNVNCPPH 360
Db 301 CSSVKYRPRYKCGSCVDGRCTPLQRTVYKMRFRCEDEGEMFSKNVMMIOCKCNVNCPPH 360
QY 361 NEASFRLYSLFNDIHKFRD 379
Db 361 NEASFRLYSLFNDIHKFRD 379

RESULT 3
US-09-348-815-2
; Sequence 2, Application US/09348815
; Patent No. 6534630

GENERAL INFORMATION:

APPLICANT: LI, HAODONG
ADAMS, MARK D
TITLE OF INVENTION: CONNECTIVE TISSUE GROWTH FACTOR-2
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: HUMAN GENOME SCIENCES, INC.
STREET: 9410 KEY WEST AVENUE
CITY: ROCKVILLE
STATE: MD
COUNTRY: US
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/348,815
FILING DATE: 08-Jul-1999
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: JONATHAN L. KLEIN
REGISTRATION NUMBER: 41,119
REFERENCE/DOCKET NUMBER: PFI26PDI
TELECOMMUNICATION INFORMATION:
TELEPHONE: 301-309-8504
TELEFAX: 301-309-8439
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 381 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-348-815-2

Query Match 92.2%; Score 1938; DB 4; Length 381;
Best Local Similarity 92.2%; Pred. No. 4,2e-170;

LENGTH: 374 AMINO ACIDS
 TYPE: AMINO ACID
 STRANDEDNESS:
 TOPOLOGY: LINEAR
 MOLECULE TYPE: PROTEIN
 US-08-468-847B-12

Query Match 86.5%; Score 1819.5; DB 1; Length 374;
 Best Local Similarity 88.3%; Pred. No. 3,9e-146;
 Matches 332; Conservative 10; Mismatches 27; Indels 7; Gaps 3;

QY 1 MSSSTRRLAVAVTLHLTLALSTCPACCHCLEAPKAPGVGLVRDGGCCCKVCAKOL 60
 DB 1 MSSRIARELALVYTLHLTRVGLSTCPADCHCLEAPKAPGVGLVRDGGCCCKVCAKOL 60
 QY 61 NECCSKTOPCDHTKGLGECNFGASTALKGICRAOSEGRPEYNSRIYONGESFOPNCKHQ 120
 DB 61 NECCSKTOPCDHTKGLGECNFGASTALKGICRAOSEGRPEYNSRIYONGESFOPNCKHQ 120
 QY 121 CTCIDGAVG-CIPLCQDELSPNLGCPNRLVYVSGCCBEWYCDDEDSIKDSDDDDL- 178
 DB 121 CTCIGMRGACIPLCQDELSPNLGCPNRLVYVSGCCBEWYCDDEDSIKDSDDDDL- 180
 QY 179 ---LGIDASEVELTRNNELIAIGKSSLRKLPVFGTEPRVLFNPLAHAGOKCIYOTTSM 235
 DB 181 GKGLGPDASEVELTRNNELIAIGKSSLRKLPVFGTEPRVLFNPLAHAGOKCIYOTTSM 238
 QY 236 QCSKSGCTGISTRTVNDNPECRVLYKETRICVEVRPCGOPYSSLSKKGCKSKTKKSPVPR 295
 DB 239 QCSKSGCTGISTRTVNDNPECRVLYKETRICVEVRPCGOPYSSLSKKGCKSKTKKSPVPR 298
 QY 296 FTYAGGSVYKRYKPKYCGSCVDGRCTPLOTRTVKKRFGCEDEGEMSKNMIQSKCNY 355
 DB 299 FTYAGGLSVYKRYKPKYCGSCVDGRCTPLOTRTVKKRFGCEDEGEMSKNMIQSKCNY 358
 QY 356 NCPHNEASPRLYSLF 371
 DB 359 NCPHNEAAPPYRLP 374

RESULT 6
 US-08-459-101A-2
 Sequence 2, Application US/08459101A
 Patent No. 5945300
 GENERAL INFORMATION:
 APPLICANT: LI, ET AL.
 TITLE OF INVENTION: Connective Tissue Growth Factor-2
 NUMBER OF SEQUENCES: 6
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
 ADDRESSEE: CECCHI, STEWART & OLSTEIN
 STREET: 6 BECKER FARM ROAD
 CITY: ROSELAND
 STATE: NEW JERSEY
 COUNTRY: USA
 ZIP: 07068
 COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5 INCH DISKETTE
 COMPUTER: IBM PS/2
 OPERATING SYSTEM: MS-DOS
 SOFTWARE: WORD PERFECT 5.1
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/459,101A
 FILING DATE: June 2, 1995
 CLASSIFICATION: 536
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: PCT/US94/07736
 FILING DATE: 12 JUL 94
 ATTORNEY/AGENT INFORMATION:
 NAME: FERRARO, GREGORY D.
 REGISTRATION NUMBER: 36,134
 REFERENCE/DOCKET NUMBER: 325800-317
 TELECOMMUNICATION INFORMATION:

TELEPHONE: 201-994-1700
 TELEFAX: 201-994-1744
 INFORMATION FOR SEQ ID NO:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 375 AMINO ACIDS
 TYPE: AMINO ACID
 STRANDEDNESS:
 TOPOLOGY: LINEAR
 MOLECULE TYPE: PROTEIN
 US-08-459-101A-2

Query Match 86.2%; Score 1812.5; DB 2; Length 375;
 Best Local Similarity 88.0%; Pred. No. 1,5e-145;
 Matches 331; Conservative 10; Mismatches 28; Indels 7; Gaps 3;

QY 1 MSSSTRRLAVAVTLHLTLALSTCPACCHCLEAPKAPGVGLVRDGGCCCKVCAKOL 60
 DB 1 MSSRIARELALVYTLHLTRVGLSTCPADCHCLEAPKAPGVGLVRDGGCCCKVCAKOL 60
 QY 61 NECCSKTOPCDHTKGLGECNFGASTALKGICRAOSEGRPEYNSRIYONGESFOPNCKHQ 120
 DB 61 NECCSKTOPCDHTKGLGECNFGASTALKGICRAOSEGRPEYNSRIYONGESFOPNCKHQ 120
 QY 121 CTCIDGAVG-CIPLCQDELSPNLGCPNRLVYVSGCCBEWYCDDEDSIKDSDDDDL- 178
 DB 121 CTCIGMRGACIPLCQDELSPNLGCPNRLVYVSGCCBEWYCDDEDSIKDSDDDDL- 180
 QY 179 ---LGIDASEVELTRNNELIAIGKSSLRKLPVFGTEPRVLFNPLAHAGOKCIYOTTSM 235
 DB 181 GKGLGPDASEVELTRNNELIAIGKSSLRKLPVFGTEPRVLFNPLAHAGOKCIYOTTSM 238
 QY 236 QCSKSGCTGISTRTVNDNPECRVLYKETRICVEVRPCGOPYSSLSKKGCKSKTKKSPVPR 295
 DB 239 QCSKSGCTGISTRTVNDNPECRVLYKETRICVEVRPCGOPYSSLSKKGCKSKTKKSPVPR 298
 QY 296 FTYAGGSVYKRYKPKYCGSCVDGRCTPLOTRTVKKRFGCEDEGEMSKNMIQSKCNY 355
 DB 299 FTYAGGLSVYKRYKPKYCGSCVDGRCTPLOTRTVKKRFGCEDEGEMSKNMIQSKCNY 358
 QY 356 NCPHNEASPRLYSLF 371
 DB 359 NCPHNEAAPPYRLP 374

RESULT 7
 US-08-468-847B-13
 Sequence 13, Application US/08468847B
 Patent No. 5780263
 GENERAL INFORMATION:
 APPLICANT: Hastings, Gregg A. and Adams, Mark D.
 TITLE OF INVENTION: Human CCN-Like Growth Factor
 NUMBER OF SEQUENCES: 20
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
 ADDRESSEE: CECCHI, STEWART & OLSTEIN
 STREET: 6 BECKER FARM ROAD
 CITY: ROSELAND
 STATE: NEW JERSEY
 COUNTRY: USA
 ZIP: 07068
 COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5 INCH DISKETTE
 COMPUTER: IBM PS/2
 OPERATING SYSTEM: MS-DOS
 SOFTWARE: WORD PERFECT 5.1
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/468,847B
 FILING DATE: 6 June 1995
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER:
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:

Query Match	45.6%;	Score 958;	DB 4;	Length 347;
Best Local Similarity	46.4%;	Pred. No. 2.6e-73;		

[illegible]

QY 178 LGLDASEVELTRNNELIAIGKSSILKRL-PVGTETPRVLFNPLHAHGQCIYOTMSQ 236
Db 169 -----KDRTAAGPALAAVRLDTEGPDPTM-----RANCLVOTTEMSA 207
QY 237 CSKSCGIGSTRVNTNDNPEGRIVKETRICVPRGQVYSSILKKGKCSKTKSPPEVRF 296
Db 208 CSKTCGIGSTRVNTNDNPEGRIVKETRICVPRGQVYSSILKKGKCSKTKSPPEVRF 267
QY 297 TYAGCSSVKKRYRPRKCGSVDRGCTPLQRTVYKMRPCEDGEMFSKNVMIOCKCNYN 356
Db 268 ELSCGTSVKTYRAKFCGCTDGRCTPHRTTLPEVERKCPDGEIMKNMFIKTCAHYN 327
QY 357 CPHNEASFRLY--SLFNDI 374
Db 328 CPDNDIEFSLYRKMYGDM 347

RESULT 10
US-08-468-847B-15
; Sequence 15, Application US/08468847B
; Patent No. 5780263
; GENERAL INFORMATION:
; APPLICANT: Hastings, Gregg A. and Adams, Mark D.
; TITLE OF INVENTION: Human CCN-like Growth Factor
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLIAN,
; ADDRESSEE: CECCHI, STEWART & OLSTEIN
; STREET: 6 BECKER FARM ROAD
; CITY: ROSELAND
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468,847B
; FILING DATE: 6 June 1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: MULLINS, J.G.
; REGISTRATION NUMBER: 33,073
; REFERENCE/DOCKET NUMBER: 325800-442
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 348 AMINO ACIDS
; TYPE: AMINO ACID
; STRANDEDNESS:
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PROTEIN
US-08-468-847B-15

Query Match 45.3%; Score 953.5; DB 1; Length 348;
Best Local Similarity 46.6%; Pred. No. 6.3e-73;
Matches 177; Conservative 61; Mismatches 103; Indels 39; Gaps 8;
QY 1 MSSSTRITLAVATLHL-TRLAL-STCPACHCPLE-APKCAPGVGLVDRGCGCCYCA 57
Db 1 MASVAGPISLAVLALCTRPATGDCSAQCQCAEAAPHCAGVSLVLDGCGCCYCA 60
QY 58 KOLNEDOSKTOPDCHTKGECNFGASSTALGICRQSEGRPEYNSRITONESPQNC 117
Db 61 KOLGELCTERPDPCPHKGLFCDFGSPANKRIGVCTAK-DSAPCVFGSGVTRSGBSFSSC 119

QY 118 KHCCTIDGAVGCTPLCPQELSLPNUGCPNPRLVKYGSCCEEMVEDSINDSLDDDD 177
Db 120 KYCCTCDGAVGCVPLCSMVRILPSPDCPPRRVKKLPBGCKCKEMVCDP----- 168
QY 178 LGLDASEVELTRNNELIAIGKSSILKRL-PVGTETPRVLFNPLHAHGQCIYOTMSQ 236
Db 169 -----KDRTAAGPALAAVRLDTEGPDPTM-----RANCLVOTTEMSA 207
QY 237 CSKSCGIGSTRVNTNDNPEGRIVKETRICVPRGQVYSSILKKGKCSKTKSPPEVRF 296
Db 208 CSKTCGIGSTRVNTNDNPEGRIVKETRICVPRGQVYSSILKKGKCSKTKSPPEVRF 267
QY 297 TYAGCSSVKKRYRPRKCGSVDRGCTPLQRTVYKMRPCEDGEMFSKNVMIOCKCNYN 356
Db 268 ELSCGTSVKTYRAKFCGCTDGRCTPHRTTLPEVERKCPDGEIMKNMFIKTCAHYN 327
QY 357 CPHNEASFRLY--SLFNDI 374
Db 328 CPDNDIEFSLYRKMYGDM 347

RESULT 11
US-09-142-569-6
; Sequence 6, Application US/09142569
; Patent No. 6413735
; GENERAL INFORMATION:
; APPLICANT: Lau, Lester F.
; TITLE OF INVENTION: Extracellular Matrix Signalling Molecules
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Muiray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; City: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/142,569
; FILING DATE: 02-Apr-1999
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Clough, David W.
; REGISTRATION NUMBER: 36,107
; REFERENCE/DOCKET NUMBER: 28758/33766
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 348 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: "f1sp12 amino acid sequence"
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-142-569-6

Query Match 45.3%; Score 953.5; DB 4; Length 348;
Best Local Similarity 46.6%; Pred. No. 6.3e-73;
Matches 177; Conservative 61; Mismatches 103; Indels 39; Gaps 8;
QY 1 MSSSTRITLAVATLHL-TRLAL-STCPACHCPLE-APKCAPGVGLVDRGCGCCYCA 57
Db 1 MASVAGPISLAVLALCTRPATGDCSAQCQCAEAAPHCAGVSLVLDGCGCCYCA 60

```

QY 58 KOLNECCKSTKOPDHHKGLKEGCFNGASSTALAGICAOSEGPCCYNSRTRONSGSPNC 117
Db 61 KQJGELICTERDPDHPKGLFCDFGSPANKRIGVCTAK-DGAPFCVGGSVTRSGESFSSC 119
QY 118 KHCCTCIDAVGCIPLCPQELSLPMIAGCPNRLKVSQCCCEWVCDEDSIKDSLDDODD 177
Db 120 KYGCTCLDGAVCVPLCSMDVRLSPDCCPFPRRYKLPKCKCKEAVCDEP----- 168
QY 178 LIGLDASEVELTRNNELLAIGKGSLLKL-PVFGTEPRVLEFNPLAHGQCIYOTTSWSQ 236
Db 169 -----KDRATVGPALAAVRLIEDTFGDPDTM-----RANCLVQTTESMA 207
QY 237 CSKSCCTGISTRTVNDNPECRLYKTRTRICEVRPCGQPIYSSLKKGKCKSTKTSPEVYR 296
Db 208 CSKTCMGISTRVTDNTDNFCRLCKCMVRPEADLEENIKKGGKCIIRTPKIAKEVVF 267
QY 297 TYVGGSSVKKYRPRKVGCSVDGRCTPLOTFTVKMRPCEGCEMFSKNVMMIOSCKNYN 356
Db 268 ELISGCTSVATRYAKRFEGVCTDGRCTPHRTTLLPVEKCPGELTKRNMNMFIKTCACHN 327
QY 357 CPAPHNEASRFLY--SLFNDI 374
Db 328 CPGDNDIFESLYRKMYGDM 347

```

RESULT 12
US-08-167-628-2

```

GENERAL INFORMATION:
APPLICANT: Grotendorst, Gary R.
APPLICANT: Bradham Jr., Douglas M.,
TITLE OF INVENTION: CONNECTIVE TISSUE GROWTH FACTOR
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Spensley Horn Jubas & Lubitz
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: US
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/167,628
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/752,427
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Wetherell, Jr. Ph.D., John W.
REGISTRATION NUMBER: 31,678
REFERENCE/DOCKET NUMBER: PD-1294
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-455-5100
TELEFAX: 619-455-5110
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 349 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
OS-08-167-628-2

```

Query Match	45.2%	Score	950;	DB 1:	Length	349;			
Best Local Similarity	45.7%	Pred. NC	1.3e-72;						
Matches	176;	Conservative	64;	Mismatches	97;	Indels	48;	Gaps	9;

```

QY 1 MSSSTRJLAVANTLHJLHRLALSTCPA-----CHCPLE-APKCAPGVLRDCCG 52
      1 : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1 MTASMGPRVAFVYL-----LALCSRPVAGQVCSRCPCDPBPAPRCAGVSLVLDCCG 56

QY 53 CKYCAKOLNEDCSKTOPCDHTKGLBECNFGASSTALKIGICRAOSRGRCPEYNSRIYONGES 112
      1 : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      57 CRYCAKOLIELCTERPDPCDHPKGLFCDFGSPANKRIIGVCTAK-DGAPCIFGTVYRSGES 115

QY 113 FQPNCKHOCTCIDGAAGCIIPLCPQELSLP.NLNGCPNRLVYKSGGCCBEMWODEDSIKDSL 172
      116 FQSSCKYQCTCIDGAAGCPLCSMVRLPSDPCPPRRVKIPLGKCCEBWVDEK----- 170

QY 173 DDDDDLLGDASEVELTRNNELIALIGKSSIKRL-IVRGTEPRVLEPNLHAHGOKIYOT 231
      171 -DQ-----TIVGPALAAARLEDTFGPPPTM-----RANCLVOT 203

QY 232 TSMQSCSKSCGIGISTRYTNDNPECRUYKETRICEVRBCGQPVYSLKKGKSKYTKSP 291
      204 TEMSASCSCKCGMGISTRYTNDAASCRLEKOSRBLCAVRCPEADLEENIKGKCKIRTPKIS 263

QY 292 EPRPRTFYAGCSSYKATTRPKYCGSSCYVDGRCTPLQIRYKAMFRCEDEGMEFSKNMMIQCSC 353
      264 KPIKFEIISGCTSMKRYTRAFKCGVCYDGRCTPHRTTLLPVEFKCPDGEVMMKNNMFIKTC 323

QY 352 KCNYNCPHPNEASFRLY--SLFNDI 374
      324 ACHYNCPCGDNIDFESLITYRKMTGDM 348
Db

```

RESULT 13
US-08-386-680-2

```

GENERAL INFORMATION:
APPLICANT: Grotendorst, Gary R.
APPLICANT: Bradham Jr., Douglas M.,
TITLE OF INVENTION: CONNECTIVE TISSUE GROWTH FACTOR
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Spensley Horn Jubas & Lublitz
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: US
ZIP: 92037

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/386,680
FILING DATE: 10-FEB-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/167,628

FILING DATE:
APPLICATION NUMBER: US/07/752,427
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Wetherell, Jr. Ph.D., John W.
REGISTRATION NUMBER: 31,678
REFERENCE/DOCKET NUMBER: PD-1294
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-455-5100
TELEFAX: 619-455-5110
INFORMATION FOR SEQ. ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 349 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-386-680-2

```

Query Match 45.2%; Score 950; DB 1; Length 349;
Best Local Similarity 45.7%; Pred. No. 1.2e-72;
Matches 176; Conservative 64; Mismatches 97; Indels 48; Gaps 9;

QY 1 MSSSTFTLAVATLHLTRALSTCPAA-----CHCLE-APKCAPGVGLVRCGCC 52
D 1 MTAASMPVAVAVL-----LALCSRPVAVGONSGPCRCDEAPRCPAGVSLVLDGCC 56
QY 53 CRYCAKOLNEDCKTOPCDHTKGLCECNFASSTALKICRAOSGRCEYNSRIYONGS 112
D 57 CRYCAKOLNEDCKTOPCDHTKGLCECNFASSTALKICRAOSGRCEYNSRIYONGS 115
QY 113 FOPNCKHOCCTIDGAVGICPLCPQELSLPILGCPNRLVKSQCCSEWVCDSDSL 172
D 116 FOSSCKRYOCTCLDGAVGCMLGSMVRLSPDCPFPRVRLPKGCCSEWVCDSEPK----- 170
QY 173 DDQDDLGLDASEVELTRNNELIAGKSSILKRL-PVFGTEPRVLPNPLAHGOKCIYOT 231
D 171 -DQ-----TVGPAALAAVRLDFTFGDPTMI-----RANCLVOT 203
QY 232 TSWSCSKSCGTGISTRTVNDNPECRVKEVTRICEVAPCGOPYSSLKGGKSKTKKSP 291
D 204 TSWSCSKSCGTGISTRTVNDNPECRVKEVTRICEVAPCGOPYSSLKGGKSKTKKSP 263
QY 292 EPPVFTYAGSSVKKRPRKYCGSCVDRCCTPLQTRTVKMRPCEDGEMFSKNVMIQSC 351
D 264 KPIKFEISGCTSKMTYRAKFCGVCIDRCCTPHRTTLVPEFKCPDGEVAKKMMFKTKC 323
QY 352 KCYNCPHPNEASFRLY--SLFNDI 374
D 324 ACHYNCBGNDIDFESLYRKMVGDM 348

RESULT 14

US-08-459-717-2
; Sequence 2, Application US/08459717
; Patent No. 5770209
; GENERAL INFORMATION:
; APPLICANT: Grotenhorst, Gary R.
; APPLICANT: Bradham Jr., Douglas M.,
; TITLE OF INVENTION: CONNECTIVE TISSUE GROWTH FACTOR
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Spensley Horn Jubas & Lubitz
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459,717
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/752,427
; FILING DATE: 30-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Wetherell, Jr. Ph.D., John W.
; REGISTRATION NUMBER: 31,678
; REFERENCE/DOCKET NUMBER: PD-1294
; TELEPHONE: 619-455-5100
; TELEFAX: 619-455-5110
; INFORMATION FOR SEO ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 349 amino acids
; TYPE: amino acid

TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-459-717-2

Query Match 45.2%; Score 950; DB 1; Length 349;
Best Local Similarity 45.7%; Pred. No. 1.2e-72;
Matches 176; Conservative 64; Mismatches 97; Indels 48; Gaps 9;

QY 1 MSSSTFTLAVATLHLTRALSTCPAA-----CHCLE-APKCAPGVGLVRCGCC 52
D 1 MTAASMPVAVAVL-----LALCSRPVAVGONSGPCRCDEAPRCPAGVSLVLDGCC 56
QY 53 CRYCAKOLNEDCKTOPCDHTKGLCECNFASSTALKICRAOSGRCEYNSRIYONGS 112
D 57 CRYCAKOLNEDCKTOPCDHTKGLCECNFASSTALKICRAOSGRCEYNSRIYONGS 115
QY 113 FOPNCKHOCCTIDGAVGICPLCPQELSLPILGCPNRLVKSQCCSEWVCDSDSL 172
D 116 FOSSCKRYOCTCLDGAVGCMLGSMVRLSPDCPFPRVRLPKGCCSEWVCDSEPK----- 170
QY 173 DDQDDLGLDASEVELTRNNELIAGKSSILKRL-PVFGTEPRVLPNPLAHGOKCIYOT 231
D 171 -DQ-----TVGPAALAAVRLDFTFGDPTMI-----RANCLVOT 203
QY 232 TSWSCSKSCGTGISTRTVNDNPECRVKEVTRICEVAPCGOPYSSLKGGKSKTKKSP 291
D 204 TSWSCSKSCGTGISTRTVNDNPECRVKEVTRICEVAPCGOPYSSLKGGKSKTKKSP 263
QY 292 EPPVFTYAGSSVKKRPRKYCGSCVDRCCTPLQTRTVKMRPCEDGEMFSKNVMIQSC 351
D 264 KPIKFEISGCTSKMTYRAKFCGVCIDRCCTPHRTTLVPEFKCPDGEVAKKMMFKTKC 323
QY 352 KCYNCPHPNEASFRLY--SLFNDI 374
D 324 ACHYNCBGNDIDFESLYRKMVGDM 348

RESULT 15

US-08-712-302-2
; Sequence 2, Application US/08712302
; Patent No. 5783187
; GENERAL INFORMATION:
; APPLICANT: Grotenhorst, Gary R.
; APPLICANT: Bradham Jr., Douglas M.,
; TITLE OF INVENTION: CONNECTIVE TISSUE GROWTH FACTOR
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Spensley Horn Jubas & Lubitz
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/712,302
; FILING DATE: 11-SEP-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/386,680
; FILING DATE: 10-FEB-1995
; APPLICATION NUMBER: US/08/167,628
; FILING DATE: 07/752,427
; ATTORNEY/AGENT INFORMATION:
; NAME: Wetherell, Jr. Ph.D., John W.
; REGISTRATION NUMBER: 31,678
; REFERENCE/DOCKET NUMBER: PD-1294

TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-455-5100
TELEFAX: 619-455-5110
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 349 amino acids
TYPE: amino acid
TOPOLOGY: 1linear
MOLECULE TYPE: protein
US-08-712-302-2

Query Match 45.2%; Score 950; DB 1; Length 349;
Best Local Similarity 45.7%; Pred. No. 1.2e-72;
Matches 176; Conservative 64; Mismatches 97; Indels 48; Gaps 9;

QY 1 MSSSTFRITLAVAVTLHLRLALSTCPAA-----CHCPLE-APKCAFGVGLVRDCCG 52
Db 1 MTAASKGPYVAFAVVL---LALCSRPAVGQNCSPRCDEPAPRCAPAGVSLVLDGCG 56
QY 53 CRYCAKQLNEDCSKTQPCDHTFKGLKCNFGASSTALKGICRAOSEGRPCENSRITYNGES 112
Db 57 CRYCAKQLGELCTERPDPCPHKGLFCDFGSPANRKIGVCTAK-DGAPCIFGGIVYRSGES 115
QY 113 FQPNCKHQCCTCIDGAVGCIPLCPQELSLPNLGCPNPRILVYSGQCCCEWVCDSDSIKDSL 172
Db 116 FQSCCKYQCTCLDGAAGCMPLCSMDVRLSPDPCPFPRVYKLPQKCCCEWVCDDEPK----- 170
QY 173 DDQDULGLDASEVELTRNNELIATGKSSLKRL-PVFTSEPRVLFNPLHAHQKCIYOT 231
Db 171 -DQ-----TVGSPALAAVRLLEDTFGPDPTM-----RANCLYOT 203
QY 232 TSMQSCSKSGTIGTSTRTVNDNPECHLYKTRICEVRCGQPYVSSLKGGKCSKTKSP 291
Db 204 TEMSASCSTCGMGISTRTVNDNASCHLEKQSRUCMVRPCPADLEENIKGKKCIRTPKIS 263
QY 292 EPVRFETYAGCSSVKKYPRYCGSCVDGRCCTPLQRTYKMRFCEDGEMFSKNVMIIQSC 351
Db 264 KPIKFELSGCTSMKTYRAKFCGVCYTDGRCCTPHRTTLLPVFPCPCPDGEVKKNMPIKTC 323
QY 352 KCNYNCPHPNEASFRLY--SLFNDI 374
Db 324 ACHYNCPGDNDFESLYRKMAYGDM 348

Search completed: August 5, 2003, 14:08:01
Job time : 14.9632 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 5, 2003, 14:06:57 ; Search time 34.4092 Seconds

(without alignments)
1308.079 Million cell updates/sec

Title: US-09-495-448a-2

Perfect score: 2103

Sequence: 1 MSSSTFRTLAVALTLHLTR.....PNEASFRLYSLFNDHKFRD 379

Scoring table:

BLAST/MSM2
Gapop 10.0 , Gapext 0.5

Searched: 451899 seqs, 118759770 residues

Total number of hits satisfying chosen parameters: 451899

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Listing first 45 summaries

Published_Applications_AA:*

1: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep:*

2: /cgn2_6/ptodata/1/pubppaa/PTCT_NEW_PUB.pep:*

3: /cgn2_6/ptodata/1/pubppaa/US06_NEW_PUB.pep:*

4: /cgn2_6/ptodata/1/pubppaa/US06_PUBCOMB.pep:*

5: /cgn2_6/ptodata/1/pubppaa/US07_NEW_PUB.pep:*

6: /cgn2_6/ptodata/1/pubppaa/PTCTUS_PUBCOMB.pep:*

7: /cgn2_6/ptodata/1/pubppaa/US08_NEW_PUB.pep:*

8: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB.pep:*

9: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep:*

10: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep:*

11: /cgn2_6/ptodata/1/pubppaa/US09C_PUBCOMB.pep:*

12: /cgn2_6/ptodata/1/pubppaa/US09_NEW_PUB.pep:*

13: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep:*

14: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep:*

15: /cgn2_6/ptodata/1/pubppaa/US10C_PUBCOMB.pep:*

16: /cgn2_6/ptodata/1/pubppaa/US10_NEW_PUB.pep:*

17: /cgn2_6/ptodata/1/pubppaa/US60_NEW_PUB.pep:*

18: /cgn2_6/ptodata/1/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2103	100.0	379	9	US-09-853-625B-11
2	2103	100.0	379	14	US-10-053-753-2
3	1938	92.2	381	11	US-09-901-910-2
4	1938	92.2	381	15	US-10-294-796-2
5	1938	92.2	455	9	US-09-925-301-1432
6	1930	91.8	381	15	US-10-205-823-84
7	1929	91.7	381	14	US-10-053-753-4
8	1819.5	86.5	374	9	US-09-853-625B-12
9	1819.5	86.5	375	11	US-09-901-910-7
10	1646.5	78.3	375	11	US-09-853-625B-13
11	960	45.6	347	15	US-10-245-977-7
12	957.5	45.5	348	14	US-10-101-040-3
13	953.5	45.3	348	9	US-09-853-625B-15
14	953.5	45.3	348	14	US-10-053-753-6
15	953.5	45.3	348	15	US-10-245-977-8

16	950	45.2	349	14	US-10-101-040-4	Sequence 4, Appl1
17	950	45.2	349	14	US-10-011-859-26	Sequence 26, Appl1
18	950	45.2	349	14	US-10-053-753-8	Sequence 8, Appl1
19	950	45.2	349	15	US-10-060-036-173	Sequence 173, Appl1
20	950	45.2	349	15	US-10-171-311-46	Sequence 78, Appl1
21	950	45.2	349	15	US-10-205-823-78	Sequence 2, Appl1
22	950	45.2	349	15	US-10-245-977-2	Sequence 2, Appl1
23	944	44.9	347	14	US-10-080-173-2	Sequence 2, Appl1
24	944	44.9	347	14	US-10-101-040-2	Sequence 2, Appl1
25	941.5	44.8	348	9	US-09-853-625B-14	Sequence 14, Appl1
26	937	44.6	349	15	US-10-245-977-6	Sequence 5, Appl1
27	919	43.7	349	15	US-10-245-977-6	Sequence 6, Appl1
28	861	40.9	351	9	US-09-853-625B-16	Sequence 16, Appl1
29	839.5	39.9	357	9	US-09-853-625B-17	Sequence 17, Appl1
30	839.5	39.9	357	14	US-10-011-859-25	Sequence 25, Appl1
31	773	36.8	345	15	US-10-112-267-3	Sequence 5, Appl1
32	773	36.8	345	15	US-10-112-267-6	Sequence 6, Appl1
33	773	36.8	345	15	US-10-001-054-50	Sequence 50, Appl1
34	773	36.8	367	15	US-10-112-267-4	Sequence 4, Appl1
35	773	36.8	367	15	US-10-112-267-8	Sequence 8, Appl1
36	772	36.7	345	15	US-10-112-267-5	Sequence 5, Appl1
37	772	36.7	345	15	US-10-112-267-21	Sequence 21, Appl1
38	772	36.7	367	15	US-10-112-267-7	Sequence 7, Appl1
39	772	36.7	367	15	US-10-112-267-22	Sequence 22, Appl1
40	758.5	36.1	345	15	US-10-112-267-11	Sequence 11, Appl1
41	758.5	36.1	367	15	US-10-112-267-12	Sequence 12, Appl1
42	622	29.6	339	15	US-10-112-267-36	Sequence 36, Appl1
43	622	29.6	354	14	US-10-011-859-2	Sequence 2, Appl1
44	622	29.6	354	14	US-10-112-267-37	Sequence 37, Appl1
45	621	29.5	339	15	US-10-112-267-32	Sequence 32, Appl1

ALIGNMENTS

RESULT 1

US-09-853-625B-11

Sequence 11, Application US/09853625B

Patent No. US20020049304A1

GENERAL INFORMATION:

APPLICANT: Haslings, Gregg A. and Adams, Mark D.

TITLE OF INVENTION: Human CCN-Like Growth Factor

NUMBER OF SEQUENCES: 20

CORRESPONDENCE ADDRESS:

ADDRESSEE: CARELLA, BYRNE, BAIN, GIULILIAN,

STREET: 6 BECKER FARM ROAD

CITY: ROSELAND

STATE: NEW JERSEY

COUNTRY: USA

ZIP: 07068

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 INCH DISKETTE

COMPUTER: IBM PS/2

OPERATING SYSTEM: MS-DOS

SOFTWARE: WORD PERFECT 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/853, 625B

FILING DATE: 14-May-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: <Unknown>

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: MULLINS, J.G.

REGISTRATION NUMBER: 33, 073

REFERENCE/DOCKET NUMBER: 325800-442

TELECOMMUNICATION INFORMATION:

TELEPHONE: 201-994-1700

TELEFAX: 201-994-1744

INFORMATION FOR SEQ ID NO: 11:

SEQUENCE CHARACTERISTICS:

LENGTH: 379 AMINO ACIDS

TYPE: AMINO ACID
STRANDEDNESS: <Unknown>
TOPOLOGY: LINEAR
MOLECULE TYPE: PROTEIN
SEQUENCE DESCRIPTION: SEQ ID NO: 11:
US-09-853-6258-11

Query Match 100.0%; Score 2103; DB 9; Length 379;
Best Local Similarity 100.0%; Pred. No. 3.2e-167;
Matches 379; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSSTFTLVAATVLLHLTRALSTCPACHCPLKAPGVGLVDDGGCCVCAKOL 60
DB 1 MSSSTFTLVAATVLLHLTRALSTCPACHCPLKAPGVGLVDDGGCCVCAKOL 60
QY 61 NEDSKTOPCDHTKGLKLECNFGASTALKGICRAOSEGRPEYNSRIYONGESFQPNCKHQ 120
DB 61 NEDSKTOPCDHTKGLKLECNFGASTALKGICRAOSEGRPEYNSRIYONGESFQPNCKHQ 120
QY 121 CTCIDGAVGCIPLCPQELSLPNIGCPNPLVKVSGCCCEWVDEDSIKDSLDDQDILLG 180
DB 121 CTCIDGAVGCIPLCPQELSLPNIGCPNPLVKVSGCCCEWVDEDSIKDSLDDQDILLG 180
QY 181 LDASEVELTRNNELIAGKSSSLKRLPVGTEPRVLFNPLHAHQKCIYQTTSMSCSKS 240
DB 181 LDASEVELTRNNELIAGKSSSLKRLPVGTEPRVLFNPLHAHQKCIYQTTSMSCSKS 240
QY 241 CGGISTRTVNDNPECLVETRICVRCGOPYSSLKKGKCKSTKKSPEVPRTYAG 300
DB 241 CGGISTRTVNDNPECLVETRICVRCGOPYSSLKKGKCKSTKKSPEVPRTYAG 300
QY 301 CSSVKYKRPKYCGSCVDGRCCTPLQRTVYKMRFRCEDEGEMFSKNVMIIOSCKCNYNCPHP 360
DB 301 CSSVKYKRPKYCGSCVDGRCCTPLQRTVYKMRFRCEDEGEMFSKNVMIIOSCKCNYNCPHP 360
QY 361 NEASFRLYSLENDIHKFRD 379
DB 361 NEASFRLYSLENDIHKFRD 379

RESULT 2

US-10-053-753-2
Sequence 2, Application US/10053753
Publication No. US20020150986A1

GENERAL INFORMATION:

APPLICANT: Lau, Lester F.
TITLE OF INVENTION: Extracellular Matrix Signalling Molecules
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/053,753
FILING DATE: 22-Jan-2002
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Clough, David W.
REGISTRATION NUMBER: 36,107
REFERENCE/DOCKET NUMBER: 28758/33766
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:
LENGTH: 379 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:

NAME/KEY: misc-feature
OTHER INFORMATION: "Mouse Cyr61 amino acid sequence"
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-053-753-2

Query Match 100.0%; Score 2103; DB 14; Length 379;
Best Local Similarity 100.0%; Pred. No. 3.2e-167;
Matches 379; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSSTFTLVAATVLLHLTRALSTCPACHCPLKAPGVGLVDDGGCCVCAKOL 60
DB 1 MSSSTFTLVAATVLLHLTRALSTCPACHCPLKAPGVGLVDDGGCCVCAKOL 60
QY 61 NEDSKTOPCDHTKGLKLECNFGASTALKGICRAOSEGRPEYNSRIYONGESFQPNCKHQ 120
DB 61 NEDSKTOPCDHTKGLKLECNFGASTALKGICRAOSEGRPEYNSRIYONGESFQPNCKHQ 120
QY 121 CTCIDGAVGCIPLCPQELSLPNIGCPNPLVKVSGCCCEWVDEDSIKDSLDDQDILLG 180
DB 121 CTCIDGAVGCIPLCPQELSLPNIGCPNPLVKVSGCCCEWVDEDSIKDSLDDQDILLG 180
QY 181 LDASEVELTRNNELIAGKSSSLKRLPVGTEPRVLFNPLHAHQKCIYQTTSMSCSKS 240
DB 181 LDASEVELTRNNELIAGKSSSLKRLPVGTEPRVLFNPLHAHQKCIYQTTSMSCSKS 240
QY 241 CGGISTRTVNDNPECLVETRICVRCGOPYSSLKKGKCKSTKKSPEVPRTYAG 300
DB 241 CGGISTRTVNDNPECLVETRICVRCGOPYSSLKKGKCKSTKKSPEVPRTYAG 300
QY 301 CSSVKYKRPKYCGSCVDGRCCTPLQRTVYKMRFRCEDEGEMFSKNVMIIOSCKCNYNCPHP 360
DB 301 CSSVKYKRPKYCGSCVDGRCCTPLQRTVYKMRFRCEDEGEMFSKNVMIIOSCKCNYNCPHP 360
QY 361 NEASFRLYSLENDIHKFRD 379
DB 361 NEASFRLYSLENDIHKFRD 379

RESULT 3

US-09-901-910-2
Sequence 2, Application US/09901910
Publication No. US20030012768A1

GENERAL INFORMATION:

APPLICANT: Li, Haodong
APPLICANT: Adams, Mark
APPLICANT: Calenda Valerie
TITLE OF INVENTION: Connective Tissue Growth Factor-2
FILE REFERENCE: PFI26P2
CURRENT APPLICATION NUMBER: US/09/901,910
CURRENT FILING DATE: 2001-07-11
PRIOR APPLICATION NUMBER: 09/348,815
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: 08/459,101
PRIOR FILING DATE: 1995-06-02
PRIOR APPLICATION NUMBER: PCT/US94/07736
PRIOR FILING DATE: 1994-07-12
PRIOR APPLICATION NUMBER: 60/217,402
PRIOR FILING DATE: 2000-07-11
PRIOR APPLICATION NUMBER: 60/291,642
PRIOR FILING DATE: 2001-05-18
NUMBER OF SEQ ID NOS: 8
SOFTWARE: PatentIn version 3.0
SEQ ID NO: 2
LENGTH: 381
TYPE: prt

ORGANISM: homo sapiens
US-09-901-910-2

Query Match 92.2%; Score 1938; DB 11; Length 381;
 Best Local Similarity 91.4%; Pred. No. 1.8e-153;
 Matches 350; Conservative 9; Mismatches 18; Indels 6; Gaps 2;

QY 1 MSSSTFTLAVAVTLHLTRIALSTCPAACHCPLEAPKCAPGVGLVDRDGGCCCKVCAKOL 60
 DB 1 MSSRIARALAVTLHLTRIALSTCPAACHCPLEAPKCAPGVGLVDRDGGCCCKVCAKOL 60

QY 61 NEDCSKTOPCDHTKGLGECNFGASSTALKGICRAOSEGRPCENSRITYONGESFOPNCKHQ 120
 DB 61 NEDCSKTOPCDHTKGLGECNFGASSTALKGICRAOSEGRPCENSRITYONGESFOPNCKHQ 120

QY 121 CTCIDGAVGCIPLCPQELSLPNLGCNPNRLVYKVGOCCEWVCDSDSIKSLDDODDL- 178
 DB 121 CTCIDGAVGCIPLCPQELSLPNLGCNPNRLVYKVGOCCEWVCDSDSIKSLDDODDLG 180

QY 179 --LGLDASEVELTRNNELIAGKSSLRPLVFGTEPRVLFPNLAHGOKCIYOTTSMQ 236
 DB 181 KELGFDASEVELTRNNELIAGKSSLRPLVFGTEPRVLFPNLAHGOKCIYOTTSMQ 238

QY 237 CSKSCGIGISTRVTNDNPECKLVETRICVRCGOPYSSSLKGGKCSKTKKSPPEYRF 296
 DB 239 CSKTCGIGISTRVTNDNPECKLVETRICVRCGOPYSSSLKGGKCSKTKKSPPEYRF 298

QY 297 TYAGCSSVKKYPKPYCGSCVDGRCTPLQRTVYKMRRCEDGEMFSKNVMMIIOCKCNYN 356
 DB 299 TYAGCLSVKKTYPKPYCGSCVDGRCTPLQRTVYKMRRCEDGEMFSKNVMMIIOCKCNYN 358

QY 357 CPHNEASFRLYSLEFNDIHKFRD 379
 DB 359 CPHANEAFPEYRLFNDIHKFRD 381

RESULT 4

US-10-294-796-2
 : Sequence 2, Application US/10294796
 : Publication No. US20030078391A1
 : GENERAL INFORMATION:
 : APPLICANT: Li, Haodong et al.
 : TITLE OF INVENTION: Connective Tissue Growth Factor-2
 : FILE REFERENCE: PFI26P1D2
 : CURRENT APPLICATION NUMBER: US/10/294,796
 : CURRENT FILING DATE: 2002-11-15
 : PRIOR APPLICATION NUMBER: US 09/348,815
 : PRIOR FILING DATE: 1999-07-08
 : PRIOR APPLICATION NUMBER: US 08/459,101
 : PRIOR FILING DATE: 1995-06-02
 : PRIOR APPLICATION NUMBER: PCT/US94/07736
 : PRIOR FILING DATE: 1994-07-12
 : NUMBER OF SEQ ID NOS: 6
 : SOFTWARE: PatentIn version 3.1
 : SEQ ID NO 2
 : LENGTH: 381
 : TYPE: PRT
 : ORGANISM: Homo sapiens
 : US-10-294-796-2

Query Match 92.2%; Score 1938; DB 15; Length 381;
 Best Local Similarity 91.4%; Pred. No. 1.8e-153;
 Matches 350; Conservative 9; Mismatches 18; Indels 6; Gaps 2;

QY 1 MSSSTFTLAVAVTLHLTRIALSTCPAACHCPLEAPKCAPGVGLVDRDGGCCCKVCAKOL 60
 DB 1 MSSRIARALAVTLHLTRIALSTCPAACHCPLEAPKCAPGVGLVDRDGGCCCKVCAKOL 60

QY 61 NEDCSKTOPCDHTKGLGECNFGASSTALKGICRAOSEGRPCENSRITYONGESFOPNCKHQ 120
 DB 61 NEDCSKTOPCDHTKGLGECNFGASSTALKGICRAOSEGRPCENSRITYONGESFOPNCKHQ 120

QY 121 CTCIDGAVGCIPLCPQELSLPNLGCNPNRLVYKVGOCCEWVCDSDSIKSLDDODDL- 178
 DB 121 CTCIDGAVGCIPLCPQELSLPNLGCNPNRLVYKVGOCCEWVCDSDSIKSLDDODDLG 180

QY 179 --LGLDASEVELTRNNELIAGKSSLRPLVFGTEPRVLFPNLAHGOKCIYOTTSMQ 236
 DB 181 KELGFDASEVELTRNNELIAGKSSLRPLVFGTEPRVLFPNLAHGOKCIYOTTSMQ 238

QY 237 CSKSCGIGISTRVTNDNPECKLVETRICVRCGOPYSSSLKGGKCSKTKKSPPEYRF 296
 DB 239 CSKTCGIGISTRVTNDNPECKLVETRICVRCGOPYSSSLKGGKCSKTKKSPPEYRF 298

QY 297 TYAGCSSVKKYPKPYCGSCVDGRCTPLQRTVYKMRRCEDGEMFSKNVMMIIOCKCNYN 356
 DB 299 TYAGCLSVKKTYPKPYCGSCVDGRCTPLQRTVYKMRRCEDGEMFSKNVMMIIOCKCNYN 358

QY 357 CPHNEASFRLYSLEFNDIHKFRD 379
 DB 359 CPHANEAFPEYRLFNDIHKFRD 381

RESULT 5

US-09-925-301-1432
 : Sequence 1432, Application US/09925301
 : Patent No. US20020052308A1
 : GENERAL INFORMATION:
 : APPLICANT: Rosen et al.
 : TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
 : FILE REFERENCE: PA106
 : CURRENT APPLICATION NUMBER: US/09/925,301
 : CURRENT FILING DATE: 2001-08-10
 : PRIOR APPLICATION NUMBER: PCT/US00/05882
 : PRIOR FILING DATE: 2000-03-08
 : PRIOR APPLICATION NUMBER: 60/124,270
 : PRIOR FILING DATE: 1999-03-12
 : NUMBER OF SEQ ID NOS: 1694
 : SOFTWARE: PatentIn Ver. 2.0
 : SEQ ID NO 1432
 : LENGTH: 455
 : TYPE: PRT
 : ORGANISM: Homo sapiens
 : US-09-925-301-1432

Query Match 92.2%; Score 1938; DB 9; Length 455;
 Best Local Similarity 91.4%; Pred. No. 2.2e-153;
 Matches 350; Conservative 9; Mismatches 18; Indels 6; Gaps 2;

QY 1 MSSSTFTLAVAVTLHLTRIALSTCPAACHCPLEAPKCAPGVGLVDRDGGCCCKVCAKOL 60
 DB 75 MSSRIARALAVTLHLTRIALSTCPAACHCPLEAPKCAPGVGLVDRDGGCCCKVCAKOL 134

QY 61 NEDCSKTOPCDHTKGLGECNFGASSTALKGICRAOSEGRPCENSRITYONGESFOPNCKHQ 120
 DB 135 NEDCSKTOPCDHTKGLGECNFGASSTALKGICRAOSEGRPCENSRITYONGESFOPNCKHQ 194

QY 121 CTCIDGAVGCIPLCPQELSLPNLGCNPNRLVYKVGOCCEWVCDSDSIKSLDDODDL- 178
 DB 195 CTCIDGAVGCIPLCPQELSLPNLGCNPNRLVYKVGOCCEWVCDSDSIKSLDDODDLG 254

QY 179 --LGLDASEVELTRNNELIAGKSSLRPLVFGTEPRVLFPNLAHGOKCIYOTTSMQ 236
 DB 255 KELGFDASEVELTRNNELIAGKSSLRPLVFGTEPRVLFPNLAHGOKCIYOTTSMQ 312

QY 237 CSKSCGIGISTRVTNDNPECKLVETRICVRCGOPYSSSLKGGKCSKTKKSPPEYRF 296
 DB 313 CSKTCGIGISTRVTNDNPECKLVETRICVRCGOPYSSSLKGGKCSKTKKSPPEYRF 372

QY 297 TYAGCSSVKKYPKPYCGSCVDGRCTPLQRTVYKMRRCEDGEMFSKNVMMIIOCKCNYN 356
 DB 373 TYAGCLSVKKTYPKPYCGSCVDGRCTPLQRTVYKMRRCEDGEMFSKNVMMIIOCKCNYN 432

QY 357 CPHNEASFRLYSLEFNDIHKFRD 379
 DB 433 CPHANEAFPEYRLFNDIHKFRD 455

```

RESULT 6
US-10-205-823-84
; Sequence 84, Application US/10205823
; Publication No. US20030108963A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Endege, Wilson O.
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Gorbacheva, Bella
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Monsey, Angela M.
; APPLICANT: Clatt, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Anderson, Dustin
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF PROSTATE CANCER
; FILE REFERENCE: MRI-044
; CURRENT APPLICATION NUMBER: US/10/205,823
; CURRENT FILING DATE: 2002-07-25
; PRIOR APPLICATION NUMBER: 60/307,982
; PRIOR FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: 60/314,356
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/325,020
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: 60/341,746
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: 60/362,158
; PRIOR FILING DATE: 2002-03-05
; NUMBER OF SEQ ID NOS: 455
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 84
; LENGTH: 381
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-205-823-84

```

Query Match	91.8%: Score 1930; DB 15; Length 381;
Best Local Similarity	90.9%: Pred. No. 8.1e-153;
Matches 348; Conservative 10; Mismatches 19; Indels 6; Gaps 2.	
QY 1.	MSSTFTFLAVAVTLLHLTRLALSTCPAACHCPLPAKCAPGVGLVDRDGGCCCKVCAKOL 60
DB 1	MSRIALALVWTLHLTRLALSTCPAACHCPLPAKCAPGVGLVDRDGGCCCKVCAKOL 60
QY 61	NEDCSKTOPCHTGTGLBCNFGASSTALKGTORASEGSPCCNYNSTIYONGSFQPNCKHQ 120
DB 61	NEDCSKTOPCHTGTGLBCNFGASSTALKGTORASEGSPCCNYNSTIYONGSFQPNCKHQ 120
QY 121	CTCIDGAVGCIPLCPQELSLPNLGCNPRLVKVSGGCCSEWVCDSDSIKDSLDDODDL - 178
DB 121	CTCIDGAVGCIPLCPQELSLPNLGCNPRLVKVSGGCCSEWVCDSDSIKDEMDODGLLG 180
QY 179	--LGLDASEVELTNNNELIAGKSSILKRLPVETEPERYVLNPLHAHOKITVOTTSKSO 236
DB 181	KELGFDASEVELTNNNELIAGKSSILKRLPVEMERIRKNPL--QOQKIVOTTSKSO 238
QY 237	CSKSCGTGISTRYTNDNDECRVLKETRICEVRCGPQVYSSLLKKKKCSKTKRSDEPVRF 296
DB 239	CSKTCGTGISTRYTNDNDECRVLKETRICEVRCGPQVYSSLLKKKKCSKTKRSDEPVRF 298
QY 297	TYAGCSSVKYKRYRPGGSCVDGRCCPTQLQFTVYVMRRCEGDEGKFSKNVMMIOSCKKNYN 356
DB 299	TYAGCSSVKYKRYRPGGSCVDGRCCPTQLQFTVYVMRRCEGDEGKFSKNVMMIOSCKKNYN 358
QY 357	CPAPNEASFRLYSLFNDIHKFRD 379
DB 359	CPHANEAAFPFYRLFNDIHKFRD 381

```

1      RESULT 7
2      US-10-053-753-4
3      ; Sequence 4, Application US/10053753
4      ; Publication No. US20020150986a1
5      ; GENERAL INFORMATION:
6      ; APPLICANT: Lau, Lester F.
7      ; TITLE OF INVENTION: Extracellular Matrix Signalling Molecules
8      ; NUMBER OF SEQUENCES: 17
9      ; CORRESPONDENCE ADDRESS:
10     ; ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun
11     ; STREET: 6300 Sears Tower, 233 South Wacker Drive
12     ; CITY: Chicago
13     ; STATE: Illinois
14     ; COUNTRY: United States of America
15     ; ZIP: 60606-6402
16     ; COMPUTER READABLE FORM:
17     ; MEDIUM TYPE: Floppy disk
18     ; COMPUTER: IBM PC compatible
19     ; OPERATING SYSTEM: PC-DOS/MS-DOS
20     ; SOFTWARE: PatentIn Release #1.0, Version #1.30
21     ; CURRENT APPLICATION DATA:
22     ; APPLICATION NUMBER: US/10/053,753
23     ; FILING DATE: 22-Jan-2002
24     ; CLASSIFICATION: <Unknown>
25     ; ATTORNEY/AGENT INFORMATION:
26     ; NAME: Clough, David W.
27     ; REGISTRATION NUMBER: 36,107
28     ; REFERENCE/DOCKET NUMBER: 28758/33766
29     ; TELECOMMUNICATION INFORMATION:
30     ; TELEPHONE: 312/474-6300
31     ; TELEFAX: 312/474-0448
32     ; TELEX: 25-3856
33     ; INFORMATION FOR SEQ ID NO: 4:
34     ; SEQUENCE CHARACTERISTICS:
35     ; LENGTH: 381 amino acids
36     ; TYPE: amino acid
37     ; TOPOLOGY: linear
38     ; MOLECULE TYPE: protein
39     ; FEATURE:
40     ; NAME/KEY: misc_feature
41     ; OTHER INFORMATION: "Human Cyt61 amino acid sequence"
42     ; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
43     US-10-053-753-4
44
45     Query Match          91.7%; Score 1929; DB 14; Length 381;
46     Best Local Similarity 90.9%; Pred. No. 9.8e-153;
47     Matches 348; Conservative 10; Mismatches 19; Indels 6; Gaps 2;
48
49     QY      1  MSSSTFRLAAVAVLLHLHTLRALSTCPAACHCPLFAPKCAAGVGLYRDGCGCCAYCAKOL 60
50     DB      1  MSSRIARLALVYVLLHLTRLALSTCPAACHCPLFAPKCAAGVGLYRDGCGCCAYCAKOL 60
51
52     QY      61  NEDSKTOPCHDTGLGLECNFGASSTALKGICRAOSEGRPCPEYNSRIYONGSEFOPNCKHO 120
53     DB      61  NEDSKTOPCHDTGLGLECNFGASSTALKGICRAOSEGRPCPEYNSRIYONGSEFOPNCKHO 120
54
55     QY      121 CTCIDGAVGCIPLCPELSTPNIGCPNRLVKVYSGQCCCEMWVCDSDISLDDQDL- 178
56     DB      121 CTCIDGAVGCIPLCPELSTPNIGCPNRLVKVYSGQCCCEMWVCDSDISLDDQDL- 180
57
58     QY      179 --LGLDASEVLTTRNNELIAAGKSSSLKRLPVFSTEPRLVLPNPLHANGOKCIYOTTSMSQ 236
59     DB      181 KELGFDASEVELTRNNELIAAGKRSLSKRLPVFEGMEPRILLYNPL--OGQKCIYOTTSMO 238
60
61     QY      237 CSKSCGAGISIRVINDNPECHLVETRIICEVRCGPQVYSSLKGGKCSKTKKSPPEVRF 296
62     DB      239 CSKTCGIGISTRVINDNPECHLVETRIICEVRCGPQGVYSSLKGGKCSKTKKSPPEVRF 298
63
64     QY      297 TYAGCSSLVKKYRPYKCGSSCYDGRCTPLQTRTYVAMRRCDEGEMFSKNVMIOSCKNCYN 356
65     DB      299 TYAGCSSLVKKYRPYKCGSSCYDGRCTPLQTRTYVAMRRCDEGEMFSKNVMIOSCKNCYN 358
66
67     YZ      357 CPHDNEASFRLYSLFNDIHKFRD 379

```


PRIOR APPLICATION NUMBER: 09/292,036
PRIOR FILING DATE: 1999-04-14
PRIOR APPLICATION NUMBER: US 09/292,036
PRIOR FILING DATE: 1999-04-14
PRIOR APPLICATION NUMBER: US 09/187,478
PRIOR FILING DATE: 1998-11-06
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn version 3.0
SEQ ID NO: 3
LENGTH: 348
TYPE: PRT
ORGANISM: Mouse
US-101-040-3

Query Match 45.5%; Score 957.5; DB 14; Length 348;
Best Local Similarity 46.8%; Pred. No. 7e-72;
Matches 178; Conservative 60; Mismatches 103; Indels 39; Gaps 8;

OY 1 MSSFTFRLAVAVTLHL-TRLAL-STCPACHCPL-APKCAPGVGLVRDGGCGCKYCA 57
DB 1 MSAVAGPISLALVILALCTRPATGDCSACQCAEAAPHCAPAGSVLYLDGCGCKYCA 60
OY 58 KOLNEDSKTOPCDHTKLECNFGASSTALKGICRAOEGRPCEVNSRIYONGESFOPNC 117
DB 61 KOLGELCTERDPCDPHKGILFCDFGSPANRKIGVCTAK-DGAPCVGGSVYRSGESFQSSC 119
OY 118 KHOCTCIDGANGCIPLCQELSLPNLGCPNRLYKVSQCCBEWVCDEDSIKSDDDDD 177
DB 120 KYQCTCLDGAAGVCLSMVDRLPSDPDPFRPVKLPKCKCEWVCDEP----- 168
OY 178 LIGLDASEVELTRNNELIAGKSSILKL-PVFGTEPVLNPLHAHQKCIYOTTSMSQ 236
DB 169 -----KDRNAVGPALAAVLEDTFGDPPTM-----RANLVOTTEMSA 207
OY 237 CSKSCGTGISTRTVNDNEPRLVKETRICEVPRGQPYSSLKKGKCKSKTKKSPYRF 296
DB 208 CSKTCGMGISTRTVNDNEPRLVKETRICEVPRGQPYSSLKKGKCKSKTKKSPYRF 267
OY 297 TYAGCSSYKRYRPRYCGSCVGRCTPIQRTVYKMRFCDEGEEMFSKVMVMIOSCKYIN 356
DB 268 ELSCGTSYKTYRAFCGCTDGRCTPHRTTLPEVEKCPDGEIMKKNMFIKTCACHYN 327
OY 357 CPHNEASFRLY--SLFNDI 374
DB 328 CPGNDIFESLYRKMYGDM 347

RESULT 13

US-09-853-625B-15
Sequence 15, Application US/09853625B
Patent No. US20020049304A1
GENERAL INFORMATION:
APPLICANT: Hastings, Gregg A. and Adams, Mark D.
TITLE OF INVENTION: Human CCN-Like Growth Factor
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
CECCHI, STEWART & OLSTEIN
STREET: 6 BECKER FARM ROAD
CITY: ROSELAND
STATE: NEW JERSEY
COUNTRY: USA
ZIP: 07068
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 INCH DISKETTE
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: WORD PERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/853,625B
FILING DATE: 14-May-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/053,587
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: WOLINS, J.G.
REGISTRATION NUMBER: 33,073
REFERENCE/DOCKET NUMBER: 325800-442
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 348 AMINO ACIDS
TYPE: AMINO ACID
STRANDEDNESS: <Unknown>
TOPOLOGY: LINEAR
MOLECULE TYPE: PROTEIN
SEQUENCE DESCRIPTION: SEQ ID NO: 15:
US-09-853-625B-15

Query Match 45.3%; Score 953.5; DB 9; Length 348;
Best Local Similarity 46.8%; Pred. No. 1.5e-71;
Matches 177; Conservative 61; Mismatches 103; Indels 39; Gaps 8;

OY 1 MSSFTFRLAVAVTLHL-TRLAL-STCPACHCPL-APKCAPGVGLVRDGGCGCKYCA 57
DB 1 MSAVAGPISLALVILALCTRPATGDCSACQCAEAAPHCAPAGSVLYLDGCGCKYCA 60
OY 58 KOLNEDSKTOPCDHTKLECNFGASSTALKGICRAOEGRPCEVNSRIYONGESFOPNC 117
DB 61 KOLGELCTERDPCDPHKGILFCDFGSPANRKIGVCTAK-DGAPCVGGSVYRSGESFQSSC 119
OY 118 KHOCTCIDGANGCIPLCQELSLPNLGCPNRLYKVSQCCBEWVCDEDSIKSDDDDD 177
DB 120 KYQCTCLDGAAGVCLSMVDRLPSDPDPFRPVKLPKCKCEWVCDEP----- 168
OY 178 LIGLDASEVELTRNNELIAGKSSILKL-PVFGTEPVLNPLHAHQKCIYOTTSMSQ 236
DB 169 -----KDRNAVGPALAAVLEDTFGDPPTM-----RANLVOTTEMSA 207
OY 237 CSKSCGTGISTRTVNDNEPRLVKETRICEVPRGQPYSSLKKGKCKSKTKKSPYRF 296
DB 208 CSKTCGMGISTRTVNDNEPRLVKETRICEVPRGQPYSSLKKGKCKSKTKKSPYRF 267
OY 297 TYAGCSSYKRYRPRYCGSCVGRCTPIQRTVYKMRFCDEGEEMFSKVMVMIOSCKYIN 356
DB 268 ELSCGTSYKTYRAFCGCTDGRCTPHRTTLPEVEKCPDGEIMKKNMFIKTCACHYN 327
OY 357 CPHNEASFRLY--SLFNDI 374
DB 328 CPGNDIFESLYRKMYGDM 347

RESULT 14

US-10-053-753-6
Sequence 6, Application US/10053753
Publication No. US20020150986A1
GENERAL INFORMATION:
APPLICANT: Lau, Lester F.
TITLE OF INVENTION: Extracellular Matrix Signalling Molecules
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/053,753
FILING DATE: 22-Jan-2002
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Clough, David W.
REGISTRATION NUMBER: 36,107
REFERENCE/DOCKET NUMBER: 28758/33766
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 348 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: misc_feature
OTHER INFORMATION: "Faspl2 amino acid sequence"
SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-10-053-753-6

Query Match 45.3%; Score 953.5; DB 14; Length 348;
Best Local Similarity 46.6%; Pred. No. 1.5e-71;
Matches 177; Conservative 61; Mismatches 103; Indels 39; Gaps 8;

QY 1 MSSSTFTLAVATLHL-TRLAL-STCPAACHPLE-APKCAPGVGLVDDGGCCCVCA 57
D 1 MLASVAGPISLAVLALCTRPATGDCSACCCAAEAAPHCAPAGSVLVLDGGCCCVCA 60
QY 58 KOLNEDCKTOPCDHTKLECNFGASSTALKGICRAOSEGRPCPEYNSRIYONGESFPQNC 117
D 61 KQJGELCTERDPCDPHKGFLCDFGSPANKRIGVCTAK-DGAPCVFGGSVRSGESFPSSC 119
QY 118 KHCCTCIDGAVGCVPLCPQELSLPNLGCNPRLVYKVGCCCEWVCDSDSIRSLDDDD 177
D 120 KYOCTCIDGAVGCVPLCSMDVRLPSPPRRVKLPKCKCKEWCDEP----- 168
QY 178 LIGLDASEVELTNNELIATGKSSSLKRL-PVFGTEPRVLFPNLHAHGQKCIYOTTSMSQ 236
D 169 -----KDRTAAGPALAAVRLDTEGPPTM-----RANCLVOTTEMSA 207
QY 237 CSKSCGTGISTRVTNDNPECRLYKETRICEVRCPOGQPVYSSLKKGKCKSTKSPPEVPF 296
D 208 CSKTCGKIGISTRVTNDNTEFCRLKQSRKLCWVRCOEADLENIKKKCIPTPIAKPVKF 267
QY 297 TYAGCSSVKYRPRKYGSCVDGRCTPLQTRIVKMRFCEDGEMFSKNVMIOSCKCNYN 356
D 268 ELGCGTSVKTYRAKFCGCTDGRCCPHRTTLLPVEFKCPDGEIMKKNMFIKTCACHYN 327
QY 357 CPHNNEASFRLY--SLFNDI 374
D 328 CPDNDIFESLYRKMTGDM 347

RESULT 15

US-10-245-977-8
Sequence 8, Application US/10245977
Publication No. US20030113816A1
GENERAL INFORMATION:
APPLICANT: Weitz, Stephen L.
APPLICANT: Usinger, William R.
TITLE OF INVENTION: METHODS OF ASSAYING CONNECTIVE TISSUE GROWTH FACTOR
FILE REFERENCE: FP0812 US
CURRENT APPLICATION NUMBER: US/10/245,977
CURRENT FILING DATE: 2002-09-18
PRIOR APPLICATION NUMBER: US 60/323,305
PRIOR FILING DATE: 2001-09-18
NUMBER OF SEQ ID NOS: 8
SOFTWARE: PatentIn version 3.1
SEQ ID NO 8

LENGTH: 348
TYPE: PRT
ORGANISM: Mus musculus
US-10-245-977-8

Query Match 45.3%; Score 953.5; DB 15; Length 348;
Best Local Similarity 46.6%; Pred. No. 1.5e-71;
Matches 177; Conservative 61; Mismatches 103; Indels 39; Gaps 8;

QY 1 MSSSTFTLAVATLHL-TRLAL-STCPAACHPLE-APKCAPGVGLVDDGGCCCVCA 57
D 1 MLASVAGPISLAVLALCTRPATGDCSACCCAAEAAPHCAPAGSVLVLDGGCCCVCA 60
QY 58 KOLNEDCKTOPCDHTKLECNFGASSTALKGICRAOSEGRPCPEYNSRIYONGESFPQNC 117
D 61 KQJGELCTERDPCDPHKGFLCDFGSPANKRIGVCTAK-DGAPCVFGGSVRSGESFPSSC 119
QY 118 KHCCTCIDGAVGCVPLCPQELSLPNLGCNPRLVYKVGCCCEWVCDSDSIRSLDDDD 177
D 120 KYOCTCIDGAVGCVPLCSMDVRLPSPPRRVKLPKCKCKEWCDEP----- 168
QY 178 LIGLDASEVELTNNELIATGKSSSLKRL-PVFGTEPRVLFPNLHAHGQKCIYOTTSMSQ 236
D 169 -----KDRTAAGPALAAVRLDTEGPPTM-----RANCLVOTTEMSA 207
QY 237 CSKSCGTGISTRVTNDNPECRLYKETRICEVRCPOGQPVYSSLKKGKCKSTKSPPEVPF 296
D 208 CSKTCGKIGISTRVTNDNTEFCRLKQSRKLCWVRCOEADLENIKKKCIPTPIAKPVKF 267
QY 297 TYAGCSSVKYRPRKYGSCVDGRCTPLQTRIVKMRFCEDGEMFSKNVMIOSCKCNYN 356
D 268 ELGCGTSVKTYRAKFCGCTDGRCCPHRTTLLPVEFKCPDGEIMKKNMFIKTCACHYN 327
QY 357 CPHNNEASFRLY--SLFNDI 374
D 328 CPDNDIFESLYRKMTGDM 347

Search completed: August 5, 2003, 14:18:50
Job time: 36.4092 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 5, 2003, 14:03:01 ; Search time 15.4592 Seconds
(without alignments)
2357.684 Million cell updates/sec

Title: US-09-495-448a-2

Sequence: 1 MSSSTFRTLAVALTLHLTR.....PNEASRLYSLEFNDIHKFRD 379

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2103	100.0	379	2	A35669	gene CYR61 protein
2	1646.5	78.3	375	2	A41428	CEF-10 protein pre
3	957.5	45.5	348	2	A40578	beta IG-M2 protein
4	950	45.2	349	2	A40551	connective tissue
5	861	40.9	351	2	S20078	NOV protein - chic
6	839.5	39.9	357	2	I38069	gene NOVH protein
7	171	8.1	1111	2	T26972	hypothetical prote
8	166	7.9	1034	2	JC5598	mucin - rat
9	157.5	7.5	13288	2	T03099	mucin, submaxillar
10	156.5	7.4	1574	2	T13954	MEGF6 protein - ra
11	156	7.4	1620	2	T27283	hypothetical prote
12	155	7.3	1042	2	A57534	secreted leucine-r
13	152.5	7.2	1056	2	A53767	mucin 5Ac (clone L
14	152	7.2	1056	2	A53767	mucin MOC5B, trach
15	151.5	7.2	1700	2	S08167	Balblani ring 3 pr
16	146	6.9	837	2	A42112	mucin-like peptide
17	145	6.9	601	2	T22025	hypothetical prote
18	145	6.9	601	2	D89711	protein F40E10.4 l
19	145	6.9	1531	2	T42218	silt-1 protein hom
20	144.5	6.9	1170	2	A53612	laminin B1k chain
21	143	6.8	1101	2	T09059	notch4 - mouse
22	142.5	6.8	1101	2	T16840	hypothetical prote
23	141	6.7	251	2	A55035	cysteine-rich prot
24	141	6.7	5376	2	T42215	zonahesin - mouse
25	140	6.7	3106	1	S53868	laminin alpha-2 ch
26	139.5	6.6	1178	1	A39604	thrombospondin pre
27	139	6.6	473	2	A56175	adhesive plaque pr
28	138.5	6.6	1847	2	T18308	probable vitelloloe
29	136.5	6.5	3020	2	A43932	mucin 2 precursor,

30	135.5	6.4	305	2	I48601	insulin-like growt
31	135	6.4	1523	2	T13953	MEGF5 protein - ra
32	135	6.4	3672	2	T23433	hypothetical prote
33	135	6.4	3704	2	T37316	probable laminin a
34	134	6.4	2219	2	T27684	hypothetical prote
35	133.5	6.3	305	2	JN0508	insulin-like growt
36	133.5	6.3	424	2	S11676	spore coat protein
37	133.5	6.3	2555	2	A40043	notch protein homo
38	133.5	6.3	3002	2	A47221	fibritillin 1 precu
39	133	6.3	1220	2	A56136	jagged protein pre
40	133	6.3	2531	2	S18188	notch protein homo
41	132.5	6.3	884	2	T18649	hypothetical prote
42	132.5	6.3	1106	2	T13938	gene shuttle craft
43	132.5	6.3	1106	2	T44598	hypothetical prote
44	132	6.3	1292	2	T09229	galactose binding
45	131.5	6.3	810	2	T10756	Nel-homology protei

ALIGNMENTS

RESULT 1

A35669

gene CYR61 protein precursor - mouse

C:Species: Mus musculus (house mouse)

C>Date: 28-Sep-1990 #sequence,revision 18-Nov-1992 #text_change 05-Nov-1999

C:Accession: A35669; I48319; S16446

R:O'Brien, T.P.; Yang, G.P.; Sanders, L.; Lau, L.F.

Mol. Cell. Biol. 10, 3569-3577, 1990

A:Title: Expression of cyr61, a growth factor-inducible immediate-early gene.

A:Reference number: A35669; MUID:90287146; PMID:2355916

A:Accession: A35669

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-379 <OAB>

A:Cross-references: GB:M32490; NID:9192909; PIDN:AAA37512.1; PID:9309206

A>Note: the authors translated the codon GAT for residue 337 as Gln

R:Latimic, B.V.; O'Brien, T.P.; Lau, L.F.

Nucleic Acids Res. 19, 3261-3267, 1991

A:Title: Promoter function and structure of the growth factor-inducible immediate ear

A:Reference number: I48319; MUID:91288203; PMID:2062642

A:Accession: I48319

A>Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-379 <RES>

A:Cross-references: EMBL:X56790; NID:950632; PIDN:CAA40109.1; PID:950633

A>Note: the authors did not translate the codon for residue 108

A>Note: the authors translated the codon GAT for residue 337 as Gln

C:Genetics:

A:Gene: CYR61

A:Insertions: 21/3; 93/1; 208/1; 279/3

C:Superfamily: von Willebrand factor type C repeat homology

F;99-166/Domain: von Willebrand factor type C repeat homology <WVC>

Query Match 100.0%; Score 2103; DB 2; Length 379;

Best Local Similarity 100.0%; Pred. No.1.5e-141;

Matches 379; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MSSSTFRTLAVALTLHLTRIALSTCPAACHCPLEAFKCAVGLVDRGCGCCVCAKOL	60
DB	1	MSSSTFRTLAVALTLHLTRIALSTCPAACHCPLEAFKCAVGLVDRGCGCCVCAKOL	60
QY	61	NEDCSKTQPCDHTKGLGECNFGASTAIGICRAOSEBPCEYNSRIYONGESFQPNCKHQ	120
DB	61	NEDCSKTQPCDHTKGLGECNFGASTAIGICRAOSEBPCEYNSRIYONGESFQPNCKHQ	120
QY	121	CTCTIDGAVGCIPLCPQELSLPNIGCPNRLVYKSGCCCEWVCDSDISKSLDDQDPLG	180
DB	121	CTCTIDGAVGCIPLCPQELSLPNIGCPNRLVYKSGCCCEWVCDSDISKSLDDQDPLG	180
QY	181	LDASEVELTRNNELIAGKSSSLKRLPVFGTEPRVLPNPLAHGQKCIYOTTSWQSKS	240
DB	181	LDASEVELTRNNELIAGKSSSLKRLPVFGTEPRVLPNPLAHGQKCIYOTTSWQSKS	240

QY 241 CGTGISTRTVNDNPECLVKEIRICEVRPCGQPVYSSLSKGGKCSKTKSPPEVFTYAG 300
 |||||
 Db 241 CGTGISTRTVNDNPECLVKEIRICEVRPCGQPVYSSLSKGGKCSKTKSPPEVFTYAG 300
 QY 301 CSSVKKYRRKYGCGSCVDGRCCTPLQTRIVKMRFRCDGEMFSKNVMIOSCKNKNCPHP 360
 |||||
 Db 301 CSSVKKYRRKYGCGSCVDGRCCTPLQTRIVKMRFRCDGEMFSKNVMIOSCKNKNCPHP 360
 QY 361 NEASFRLYSLENDIHKFRD 379
 |||||
 Db 361 NEASFRLYSLENDIHKFRD 379

RESULT 2

AA1428
 CEF-10 protein precursor - chicken

C:Species: Gallus gallus (chicken)
 C:date: 03-Apr-1992 #sequence_revision 03-Apr-1992 #text_change 21-Jul-2000
 C:Accession: A41428

R:Simmons, D.L.; Levy, D.B.; Yannoni, Y.; Erikson, R.L.
 Proc. Natl. Acad. Sci. U.S.A. 86, 1178-1182, 1989

A:title: Identification of a phorbol ester-repressible v-src-inducible gene.
 A:Reference number: A41428; MUID:89145206; PMID:2537491

A:Accession: A41428
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-375 <SIM>

A:Cross-references: GB:J04496; NID:9211435; PIDN:AAA8661.1; PID:9211436

Query Match 78.3%; Score 1646.5; DB 2; Length 375;
 Best Local Similarity 80.1%; Pred. No. 2.9e-109;
 Matches 302; Conservative 22; Mismatches 36; Indels 17; Gaps 7;

QY 10 AAVATLHLTRIAL-STCPAACHPLEAPKCAPGVGLVDDGGCCGVCAKOLNEDCSKRO 68
 ||:|:|||||
 Db 9 ALAAALLCARLALSPCPAVCCPAAPQACAGVGLVDDGGCCGVCAKOLNEDCSKRO 68
 QY 69 PCDHTKGLCECNFGASSTALKGICRAOSEGRPCENSRIVONGESFOQNCNHOCTCIDGAV 128
 |||||
 Db 69 PCDHTKGLCECNFGASSTALKGICRAOSEGRPCENSRIVONGESFOQNCNHOCTCIDGAV 128
 QY 129 GCTPLCPQELSLPNCPCPNRLVYKVSGCCCEEWVCEDESIKOSLDQDDL---LGLDAS 184
 |||||
 Db 129 GCTPLCPQELSLPNCPCPNRLVYKVSGCCCEEWVCEDESIKOSLDQDDL---LGLDAS 184
 QY 185 EVELFTNNELIATIGKSSLSKRLPVFETEP--RYLNPPLAHGOKICIVOTTSNOSGSKSG 242
 |||||
 Db 187 EVELFTNNELIATIGKSSLSKRLPVFETEP--RYLNPPLAHGOKICIVOTTSNOSGSKSG 242
 QY 243 TGISTRVTNDNPECLVKEIRICEVRPCGQPVYSSLSKGGKCSKTKSPPEVFTYAGCS 302
 |||||
 Db 240 TGISTRVTNDNPECLVKEIRICEVRPCGQPVYSSLSKGGKCSKTKSPPEVFTYAGCS 299
 QY 303 SVKKYRRKYGCGSCVDGRCCTPLQTRIVKMRFRCDGEMFSKNVMIOSCKNKNCPHP 362
 |||||
 Db 300 SVKKYRRKYGCGSCVDGRCCTPLQTRIVKMRFRCDGEMFSKNVMIOSCKNKNCPHP 359
 QY 363 ASFRLYSLENDIHKFRD 379
 ||:|:|||||
 Db 360 ASFRLYSLENDIHKFRD 375

RESULT 3

AA0578
 beta IG-M2 protein precursor - mouse

C:Species: Mus musculus (house mouse)
 C:date: 06-Mar-1992 #sequence_revision 06-Mar-1992 #text_change 01-Dec-2000
 C:Accession: A40578; A53228

R:Brummer, A.; Chinn, J.; Neubauer, M.; Purchio, A.F.
 DNA Cell Biol. 10, 293-300, 1991

A:title: Identification of a gene family regulated by transforming growth factor-beta.
 A:Reference number: A40578; MUID:91229699; PMID:2029337

A:Accession: A40578
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-348

 A:Cross-references: GB:M80263; NID:9201945; PIDN:AAA73135.1; PID:9201946
 R:Hysek, R.P.; Macdonald-Bravo, H.; Mattei, M.G.; Bravo, R.
 Cell Growth Differ. 2, 225-233, 1991
 A:title: Structure, mapping, and expression of flsp-12, a growth factor-inducible gene
 A:Reference number: A53228; MUID:91363290; PMID:1888698
 A:Accession: A53228
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-160, 'K', 162-348 <RYS>
 A:Cross-references: GB:M70641; NID:9193313; PIDN:AAA37627.1; PID:9193314
 C:Genetics:
 A:Gene: flsp-12

Query Match 45.5%; Score 957.5; DB 2; Length 348;
 Best Local Similarity 46.8%; Pred. No. 1.4e-60;
 Matches 178; Conservative 60; Mismatches 103; Indels 39; Gaps 8;

QY 1 MSSSTRFLAVATLHLTRIAL-STCPAACHPLE-APKCAPGVGLVDDGGCCGVCAK 57
 ||:|:|||||
 Db 1 MLASVAGPISLALVLLALCTRRATGDCSAQCQCAEAAPHPACVSLVDDCGCCRCVCA 60
 QY 58 KOLNEDCSKTPQCDHTKGLCECNFGASSTALKGICRAOSEGRPCENSRIVONGESFOQNC 117
 |||||
 Db 61 KOLGELCTERDPCDHPKGLFCDFGSPANKRIGVCTAK-DGACVFGGSVYRSGESFOSSC 119
 QY 118 KQCTCIDGAVGCTPLCPQELSLPNCPCPNRLVYKVSGCCCEEWVCEDESIKOSLDQDDL 177
 |||||
 Db 120 KQCTCIDGAVGCTPLCPQELSLPNCPCPNRLVYKVSGCCCEEWVCEDESIKOSLDQDDL 168
 QY 178 LGLDASEVELFTNNELIATIGKSSLSKRLPVFETEP--RYLNPPLAHGOKICIVOTTSNOSQ 236
 ||:|:|||||
 Db 169 LGLDASEVELFTNNELIATIGKSSLSKRLPVFETEP--RYLNPPLAHGOKICIVOTTSNOSQ 207
 QY 237 CSKSGGTGISTRTVNDNPECLVKEIRICEVRPCGQPVYSSLSKGGKCSKTKSPPEVFTY 296
 |||||
 Db 208 CSKSGGTGISTRTVNDNPECLVKEIRICEVRPCGQPVYSSLSKGGKCSKTKSPPEVFTY 267
 QY 297 TYAGSSVKKYRRKYGCGSCVDGRCCTPLQTRIVKMRFRCDGEMFSKNVMIOSCKNKNCPHP 356
 |||||
 Db 268 ELISGCTSVKTYAKRCGVCTDRCCTPHRTTLTPVEFKCPDGLIMKMMFKTCACHYN 327
 QY 357 CPHPNEASFRLYSLENDIHKFRD 374
 ||:|:|||||
 Db 328 CPHPNEASFRLYSLENDIHKFRD 347

RESULT 4

AA0551
 connective tissue growth factor - human

C:Species: Homo sapiens (man)
 C:date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 21-Jul-2000
 C:Accession: A40551; S44205

R:Bradham, D.M.; Igarashi, A.; Potter, R.L.; Grotenhorst, G.R.
 J. Cell Biol. 114, 1285-1294, 1991

A:title: Connective tissue growth factor: a cysteine-rich mitogen secreted by human v
 A:Reference number: A40551; MUID:91373462; PMID:1654338

A:Accession: A40551
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-349

A:Cross-references: GB:M92934; GB:M36965; GB:S56201; NID:9180923; PIDN:AAA91279.1; PI
 R:Oemar, B.S.; Werner, A.; Yang, Z.; Garner, J.M.; Gentz, R.; Luescher, T.F.

submitted to the EMBL Data Library, April 1994
 A:Description: Differential cloning and expression of human connective tissue growth

A:Reference number: S44205
 A:Accession: S44205
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-349 <OEM>

A:Cross-references: EMBL:X78947; NID:9474933; PID:9474934

submitted to the EMBL Data Library, October 1998

A:Reference number: Z20293

A:Accession: T26972

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1111 <MIL>

A:Cross-references: EMBL:AL032657, PIDN:CAA21739.1, GSPDB:GN00019, CESP:Y47H9C.4

A:Experimental source: clone Y47H9C

C:Genetics:

A:Gene: CESP:Y47H9C.4

A:Map position: 1

A:Insertions: 50/2; 84/2; 150/1; 238/3; 342/3; 797/1; 851/1; 947/2; 1017/1; 1083/1

C:Superfamily: unassigned ankyrin repeat proteins; ankyrin repeat homology; EGF homology

Query Match 8.1%; Score 171; DB 2; Length 1111;
Best Local Similarity 20.1%; Pred. No. 0.00016;
Matches 86; Conservative 32; Mismatches 156; Indels 154; Gaps 20;

```

QY 26 CPACACHPLE-ARKCAGVGLV-----DGGCCCKVCAK-QLNEDSKTOPCHTKGLEC 78
DB 380 CSRTCTCVRNTLMCAVNTFCRCRKGPFYGDNCCLACSKDSYGPNCCKAMCMNNASEC 439
QY 79 NFASSTALGICRAOSEGRPCENSRITYNGSEFOPCKHOCCT-----IDGA----- 127
DB 440 NPETGSC-----VCKPRTGKNCSEPCPL-----DFYGPNCNAHQCCQCNORVGCDBADGCKQ 491
QY 128 -----VGCIPD-----CPQELSLPN--LGCP-- 146
DB 492 CDKGWTHGRCEHHCPADTFGANCEKRCRCKPGIGCDPIITGECPCPAGLQCANCDIGCPDG 551
QY 147 -----NPLRYKVSQGCCCEWVDESDISLDDODDLGLDASEVELRNNE 193
DB 552 STYRPGCKLHCKCVNGCKDKETGEC---TC-----OQCFEGSDCS----- 587
QY 194 LIAIGKSSSLKRLPVETGEERVLFPNPLAHGOKCIYO-TTWSQCSK---SCGTGISTRY 249
DB 588 -TTSCSK-----YGESELCSPCSDASCSQTKCLCPLCTKG 625
QY 250 TNDNPECRIVKTRICEVRPCGQPVYSSLRKGRKCSKTKRSPPEVFTY-----AGC 301
DB 626 VSCDQCKDPMTPFGELTCEFTVTPSPCASTDPKNGVCLSCPGSSGSIHCENHCPAGSYGDGC 685
QY 302 SSYKKYRPRKCGSCVGRCTPIQTRIV-----KMRRCEDEGEFMSKNVMIQSKC 353
DB 686 QQV-----C-SCADGHGDDPTTGECICEPGYHGKTCSKCDPKGYGCALDPCPKAS 737
QY 354 NYNCPHPN 361
DB 738 GSWCDHIN 745

```

RESULT 8

JC5598

mucin - rat

C:Species: Rattus norvegicus (Norway rat)

G:Date: 23-Sep-1997 #sequence_revision 23-Sep-1997 #text_change 05-Nov-1999

C:Accession: JC5598

R:Inatomi, T.; Tisdale, A.S.; Zhan, Q.; Spurr-Michaud, S.; Gipson, I.K.

Biochem. Biophys. Res. Commun. 236, 789-797, 1997

A:Title: Cloning of rat Muc5AC mucin gene: Comparison of its structure and tissue distri-

A:Reference number: JC5598; MUID:97396181; PMID:9245735

A:Accession: JC5598

A:Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-1034 <INA>

A:Cross-references: GB:U83139; NID:q2315984; PIDN:AAC53312.1; PID:q2315985

C:Note: translation not complete

C:Comment: This protein is a high molecular weight glycoprotein which is a major compone

ntestinal tract and reproductive tract.

C:Genetics:

A:Gene: Muc5A

A:Superfamily: von Willebrand factor type C repeat homology

F:45-149/Domain: cysteine-rich <CTS>

F:762-830/Domain: von Willebrand factor type C repeat homology <VWC>

Query Match

Best Local Similarity 21.4%; Score 166; DB 2; Length 1034;

Matches 99; Conservative 47; Mismatches 170; Indels 146; Gaps 25;

```

QY 30 CHCPLEAPKCAPGVGLVRDGGCCCKVC---AKQLNEDSKT-----OPCHTKGLECN 79
DB 558 CHM-LDLEVYCSGLIELYASLCAAGVCIPMRSHNTNMTCPFTCPENOVOPCGSPNHYCY 616
QY 80 FGASSTALGICRA--QSEBRPCENSRITY-----NESPDP----- 115
DB 617 RNDDISLALIQAGPKSECCFCPDMTLESSNDSCVPSQMCCLCPHDEPVEPGHTISI 676
QY 116 NCKHOCCTIDGAVGC-IPLCPOELSLPNLCPNPLRVKVS-----GCCCEWVDESDSK 169
DB 677 NCQ-DCICKRGITLCEKLCQPT-----CPRGVPVPSIALNAQCCSQSFSC-----VC 725
QY 170 DSLDDDDDLGLDASEVELTRNNELIALIGK-SLKLRLPVFTGEPVRLNP-----LHAG 224
DB 726 NSHCHCPPPLHCPSSSLIVYEGTCCPQSCSSQKGCVDNGT---LYOPGDVSSSLC 781
QY 225 QKCIYQVTS-----WSQCSKSGTGSTVYT-----NNPPEC 256
DB 782 ERLICEVSSNAESDVAVVNCETELCMTQCPKGFERYOTTGHCQGCQCVPTCPKKNNST 841
QY 257 RLVEK-----TRICE-----VR-----PCGQP-- 279
DB 842 SLYKPGFEMFEPNCPVTHNCEKFDVLPVAVVTKIECPKINCPDMAQLREDGCDYCLV 901
QY 280 KKKKSKTKRSPPEVFTTAAGCSSVKKRYKVC-GSCVDS-----RCQIP 323
DB 902 POKCKTVHQRQ-QILR-QQNCSESPVSLSTYQGNCGSTSMYSLEANTVEHTCECOE 958
QY 324 LQRTVKMRPCEDGEFMSKNVMIQSKC-NYNCPHPNAS 364
DB 959 LQTSQSRVTLHCDGSSRFSTYQVEKCGCLGRCRAPDTS 1000

```

RESULT 9

T03099

mucin, submaxillary - pig

N:Alternate names: apomucin

C:Species: Sus scrofa domestica (domestic pig)

G:Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 02-Jun-2000

C:Accession: T03099; A40009; A28528; B29789

R:Reckhardt, A.E.; Timpe, C.S.; Deluca, A.W.; Hill, R.L.

J. Biol. Chem. 272, 33204-33210, 1997

A:Title: The complete cDNA sequence and structural polymorphism of the polypeptide ch

A:Reference number: Z14839; MUID:98070526; PMID:9407109

A:Accession: T03099

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1328 <ECK>

A:Cross-references: EMBL:AF005273; NID:92581863; PIDN:AA62527.1; PID:q2581864

R:Reckhardt, A.E.; Timpe, C.S.; Abernethy, J.L.; Zhao, Y.; Hill, R.L.

J. Biol. Chem. 266, 9678-9686, 1991

A:Title: Porcine submaxillary mucin contains a cysteine-rich, carboxyl-terminal domain

A:Reference number: A40009; MUID:91236743; PMID:2033060

A:Accession: A40009

A:Molecule type: mRNA

A:Residues: 12139-12167, 'r', 12169-13288 <EC3>

A:Cross-references: GB:M61883; NID:q454837; PIDN:AAA30998.1; PID:q164374

R:Timpe, C.S.; Reckhardt, A.E.; Abernethy, J.L.; Hill, R.L.

J. Biol. Chem. 263, 1081-1088, 1988

A:Title: Porcine submaxillary gland apomucin contains tandemly repeated, identical se

A:Reference number: A28528; MUID:88087170; PMID:2826455

A:Accession: A28528

A:Molecule type: mRNA

A:Residues: 12139-12167, 'r', 12169-12641 <TIM>

A:Cross-references: GB:M21174; GB:J03512; NID:g164321; PIDN:AAA30990.1; PID:g552360

A:Experimental source: submaxillary gland

R:Reckhardt, A.E.; Timpe, C.S.; Abernethy, J.L.; Tounadje, A.; Johnson Jr., W.C.; Hill

J. Biol. Chem. 262, 11339-11344, 1987
 A:Title: Structural properties of porcine submaxillary gland apomucin.
 A:Reference number: A92606; MUID:87280230; PMID:3611111
 A:Accession: B29789
 A:Molecule type: protein
 A:Residues: 1572-1607 <EC2>
 C:Superfamily: pig submaxillary mucin
 C:Keywords: tandem repeat

Query Match 7.5%; Score 157.5; DB 2; Length 13288;
 Best Local Similarity 20.1%; Pred. No. 0.013;
 Matches 93; Conservative 46; Mismatches 148; Indels 175; Gaps 22;

OY 26 CPACHCPLEAPKAPGVGLVRDGGCCCKVCAKQLENECSKTQPCDHTKGLNECNAGAST 85
 DB 862 CKRCQYCP-----VGMVRSKGNK-----VFPEDC----- 886
 OY 86 ALKGICRAOSEGRPEYNSRIYONGESFQPNCKHOCTCIDGAVGCP----- 132
 DB 887 -----PCSFGRREYDQGSVTSVGC-NKCTCIKSMNCTONECQTTCHYEG 932
 OY 133 -----LCPQLSLPNIGCPN--RLVKSGGCCCEWVDCDEDSIKSLDDQD 176
 DB 933 HIRFDGKTYFSDGLQYSLFDYCGSENGTFRILTESVPCCEDELTCRKRIYAFQDON 992
 OY 177 DLL-----GLDASEVELRNNEEL-AIGKSSLRPLVF-----GTEPRYLEMPL 220
 DB 993 VLIHDKVTAVKTESKECELENSYVHTVGLYLILFLSGITIIIMDKNRISVILDP- 1051
 OY 221 HAHGKRC-----IVQTTSMSCSKSGTGISTRTVNDNPE 255
 DB 1052 RMNGVCGLCGNNNDLKDPTTRSSVAAGLENGNWKTSQESDPTVTSQFPCDSNPY 1111
 OY 256 CRVKEIRICEY-----RPGQPYYS-----LKKGKSKTKKSEPVRYTYAG-CS 302
 DB 1112 CK-AWAVRKCCEIIRDSTFRDCHNKVDPSAYDACIEEACADMEEK-----YLGFT 1162
 OY 303 SVKKY-----RKYCGSCVD-----GRCC--TPLOTRTVKMFRCEDEM 340
 DB 1163 AVAMIAEACSAVAGCVTRMKDPLCPVCDYNNAPBECSMRYEPCTVYAK--TCKDRVI 1219
 OY 341 FSKNYMMIOQC--KCNYNCPHNEASFRLYSL-----FNDI 374
 DB 1220 GQKFSALLEGCAKCPDSAPYIDENMTMKCVSLSECSCTYNDI 1261

RESULT 10
 T13954

MEGF6 protein - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jul-2000
 C:Accession: T13954
 R:Nakayama, M.; Nakajima, D.; Nagase, T.; Nomura, N.; Seki, N.; Ohara, O.
 Genomics 51, 27-34, 1998
 A:Title: Identification of high-molecular-weight proteins with multiple EGF-like motifs
 A:Reference number: Z14126; MUID:96360089; PMID:9692030
 A:Accession: T13954
 A:Status: preliminary; translated from GR/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-1574 <NAK>
 A:Cross-references: EMBL:AB011532; NID:93449293; PIDN:BA43462.1; PID:93449294
 A:Experimental source: Strain Sprague-Dawley; brain
 C:Genetics: MEGF6
 A:Gene: MEGF6

Query Match 7.4%; Score 156.5; DB 2; Length 1574;
 Best Local Similarity 20.4%; Pred. No. 0.0024;
 Matches 94; Conservative 40; Mismatches 142; Indels 185; Gaps 24;

OY 26 CPACHCP-----LEAPKAPGV--GLVRDGC----- 50
 DB 568 CSSPCTCQNGTCDPVLGACRCRCPGVSAHCEGCPKGFYGHKHKCHCANRGCHRLY 627

OY 51 GCKRVCAKOL-----NEDCSKTQPCDHTKGLNECNAGASTALKGICR 92
 DB 628 GAC-LCDGGLGRFCHLACPRAPBPGSCSEDCLEQ--SHTRSCNPKGSGS-----CK 678
 OY 93 AOSEGRPEYNSRIYONGESFQPNCKHOCTCIDGAVGCIPL-----CPQ 136
 DB 679 AGFQGERQAECE--ESG-FRPGGRCHHCTCPG--VACDPVSGECRTGCPQYGEDCGQ 733
 OY 137 ELISLNLG-----CENPRLYKVSQC--CEWVDCDEDSIKSLDDQDGLGLDASEYE 187
 DB 734 ECPVGTFFGVNCGSGSCVAGAPCHPRTVGBCLCPKGTGDBD--GADCEGRWGLCOEET- 789
 OY 188 LTRNNELIAGKSSLRPLVFTGTEPRVLENLHAHGKCIYQTS-----WSQSK 239
 DB 790 -----CPACHGASCPNETGTICLIRFVSGKROD 819
 OY 240 SC-----GTGISTR--VTNDN-----PCRLVKEIRICE-----VRPC-- 270
 DB 820 TCSAGWGTGQIRCAACANDGHCDPTTGRCSAPGMTGLSCGRACDSGHMGPDCIHPNC 879
 OY 271 -----GQPYSSLKKGKSKTKKSEPVRYTYAGSSVKKTRPKYCGSC--VD 317
 DB 880 SAGHCNDAVSGILCEAGYEGPRCEQS-----CROGYYPGSCDQKRCRCEHGAACDHVS 933
 OY 318 GRCCPLOTRTVKMFRCEDEMFSKNYMMIOQSKCNCP 358
 DB 934 GACTCPAGMRGSPCEHACPAGEF--GLDCDSACNCSAGAP 971

RESULT 11
 T27283

hypothetical protein Y64G10A.f - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T27283
 R:Ainscough, R.
 submitted to the EMBL Data Library, September 1999
 A:Reference number: Z20336
 A:Accession: T27283
 A:Status: preliminary; translated from GR/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1620 <WIL>
 A:Cross-references: EMBL:AL110498; NID:e1542303; PIDN:CA54471.1; CESP:Y64G10A.f
 C:Genetics: C:Genetics: CESP:Y64G10A.f
 A:Introns: 77/1; 116/1; 198/1; 282/1; 365/1; 425/1; 466/1; 548/1; 559/1; 601/1; 625/1

Query Match 7.4%; Score 156.5; DB 2; Length 1620;
 Best Local Similarity 23.2%; Pred. No. 0.0024;
 Matches 92; Conservative 27; Mismatches 134; Indels 143; Gaps 22;

OY 26 CPACHCPLEAP-----KCAPG-----VLVRDGGCCCKVCAKQLENECSK 66
 DB 1052 CKGICSCQNGATCDSDVTSCECRPRMRRKKCDRCRPPD--KFGECNMIICDCTTNDISM 1109
 OY 67 TOP-----CDHTKGLNECNFAGASTALKGICRAOSEGRPEYNSRIYONGESFQPNCKHOCT 122
 DB 1110 YNPEVARCDHYTG-ECR-----CPAGMTGPDQGTSCPLGRHGE--GCRHSCQ 1153
 OY 123 CIDGA-----VGCIPRLCPQLSLPNIGCPNRLVKSGGCCCEWVDCDE 165
 DB 1154 CSNGASCDRVYGFCDCCPGEKGNKNESECPGLWGSN--CMKHCLCMHGECNKE----- 1206
 OY 166 DSIKSLDDQDGLGLDASEVELTRNNELIAGKSSLRPLVFTGTEPRVLENLHAHG 225
 DB 1207 -----NEDGCIIDGMPGSPSLCPRGGRGSRNCAQCN--CKNGA 1241
 OY 226 KCIYOT-----TSNS--QCSKSGCTGISTRTVNDNPCRILVKEIRICEVRCQPYYS 277
 DB 1242 SCDRRTGRCECLPFGSGHCEKSCVSG-----HYGAKC--EFTCEBENALDPIIGH 1292
 OY 278 LK-----KGKSKTKKSKSEPVRYTYAGSSVKKTRPKYCGSC--VDGRC----- 320

Db 1293 CSCOPGWRGKCKM-----PCLKGYGRHCSCRCRANSKSCDHISGRCCPKGYAGHS 1346
 QY 321 CTPLOTRTVMKRFCEDEGEKSKVMVMIOSCKKYN 356
 Db 1347 CTEL-----CPDG-TFGSCS--OKCDDGEN 1369

RESULT 12

T42626

secreted leucine-rich repeat-containing protein slt2 - mouse (fragment)

N:Alternate names: neurogenic extracellular slt protein

C:Species: Mus musculus (house mouse)

C>Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 16-Aug-2002

C:Accession: T42626

R:Holmes, G.P.; Negus, K.; BurrIDGE, L.; Raman, S.; Alagar, E.; Yamada, T.; Little, M.H.

Mech. Dev. 79, 57-72, 1998

A:Title: Distinct but overlapping expression patterns of two vertebrate slt homologs in

A:Reference number: 22177; MUID:9279238; PMID:10349621

A:Accession: T42626

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1025 <HOL>

A:Cross-references: EMBL:AF074960; NID:g4151258; PID:g4151259; PIDN:AD04345.1

C:Genetics:

A:Gene: slt2

C:Superfamily: fruit fly slt protein; EGF homology; leucine-rich alpha-2-glycoprotein

Query Match 7.48; Score 156; DB 2; Length 1025;

Best Local Similarity 18.58; Pred. No. 0.0018;

Matches 101; Conservative 41; Mismatches 150; Indels 248; Gaps 23;

QY 24 STC-----PACHCPLPAPKCAPGVGLVRDGGCGCKYCAQOLNEDCSKTOPCDH----- 72
 Db 507 STCVGGINNYTCLCPPEYTG-----ELCEKELDFCAQDLNCPQIDSKCIL 551
 QY 73 -TGLECNFGASSALKGICRAOSEGRPCRYNSRIYNGSFQPNCHQCTCIDGANGCI 131
 Db 552 TPFGFKD-----CTPGYIGHCIDIDPDCCDNK-----CKNGAHCDDAVNGYT 595
 QY 132 PLCPQ-----ELSLP-----NLGCPNPR--LVKYS-----GOCCEW 161
 Db 596 CVCPEGSGLCFESPPMVLPRTPSPCNFPCONCAOCTIRINETICCLPGYLGEKCKL 655
 QY 162 VCEDESLKDS-----LDQDDLLGL-----DASEVELTRNNELIAI 197
 Db 656 VSNFVNRKESYLQIPSAKVRPQNTITLQIATDEDSGLLYKGDKDHIAVELYGRVASY 715
 QY 198 GKRS-----SLKRLPYGTEPRVLEN-----PL 220
 Db 716 DTGSHPASALYSVETINDGNFHVLELLTLDLSLSLVYDGGSPRYITNLKOSTLNEDSPL 775
 QY 221 HA-----HG-----OKTIYQ 230
 Db 776 YVGMPEKNVASTLRAPGONGTSFHCITRNLYINSELQDFRKMPQGTGLPCEPCCHK 835
 QY 231 TTSMQCSKSGTGISTRV-----TND----- 252
 Db 836 VCAHGMQOPSSGGECEEGMMGPLCDORTNDPCLGKNCVHGTCPLINAFYSCKCLE 895
 QY 253 -----NPECRLVK--ETRLCEVRPCGQPV--YSLAKGKCKCK--TKKSEPV 294
 Db 896 GHGCVLDEDEDLFNP--COMIKCKHGKCRSLGVGPCECNSGFTGSDSCREISCRERI 954
 QY 295 R-----FTYAGCSSVKKYKPYKC--GSCVDGRCTPLQTRTVKMRFCEDGEMSKVM 347
 Db 955 RQYKQKQGYAAACOTTKKVRLECRGCGAGGCGCCPLRSKRRTYSPECTDGSFVDEVER 1014
 QY 348 IQSCCK 353
 Db 1015 VVKCGC 1020

RESULT 13

A57534

mucin 5AC (clone L31) - human (fragment)

C:Species: Homo sapiens (man)

C>Date: 08-Feb-1996 #sequence_revision 08-Feb-1996 #text_change 20-Apr-2000

C:Accession: A57534

R:Lesuffleur, T.; Roche, F.; Hill, A.S.; Lacasa, M.; Fox, M.; Swallow, D.M.; Zweibaum

J. Biol. Chem. 270, 13665-13673, 1995

A:Title: Characterization of a mucin cDNA clone isolated from HT-29 mucous-secreting c

A:Reference number: A57534; MUID:95293957; PMID:7775418

A:Accession: A57534

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-1042 <LES>

A:Cross-references: GB:248314; NID:g1052607; PIDN:CAA88307.1; PID:g1052608

C:Genetics:

A:Gene: GDB:MUC5AC

A:Cross-references: GDB:454136; OMIM:158373

A:Map position: 11p15.5-1p15.5

A:Superfamily: von Willebrand factor type C repeat homology

F:678-746/Domain: von Willebrand factor type C repeat homology <WMC>

Query Match 7.38; Score 152.5; DB 2; Length 1042;

Best Local Similarity 18.98; Pred. No. 0.0037;

Matches 105; Conservative 48; Mismatches 159; Indels 245; Gaps 27;

QY 4 STFRITAVAVTLHLRLALSTCPACHCPLPAPKCAPGVGLVRDGC-----GC 52
 Db 411 TTVGSGTVPPTVGTGFTVPAPC---LPSPICHLISLVFECCHVYIPPLFYECC 467
 QY 53 -----CKYCAK-----QLNECSKT-----QCCHITKGL 76
 Db 468 VEDRCHMTDLDVYSSLELYAALCASHDICIDNRGRTGHMCPTCPADKRVYQPCGSNPS 527
 QY 77 ECNFGASPTALKGICRAQ--SEGRPCRYNSRIYON-----GE 111
 Db 528 YC-YGNDASLGLAPRPGFTTEGCFPEEGITLSTNAQCVPTGCRCLGPHGEYKVG 586
 QY 112 SFQPNCHQCTC--IDAVGCLP-LCPQELSLPNLCCPNRLVKV-----SGCCCEWVC 163
 Db 587 TVGMDQQ--ECTCPAATWTILTCRPLCP---LPP-ACPLPGFVVPAPAPQAGCCPQYSC 640
 QY 164 DEDSIDSLDDQDDLLGLDASEVELTRNNELIAGGSSISLKLPPV-FGEPRRLFNLPLA 222
 Db 641 -----ACNTRKCPAVGCEGARALPTQ 664
 QY 223 HGCKIYQTSWQCSKSGCTGISTRYTNDNPECRLVK-----ETRICE 266
 Db 665 EGACCPYQNCSTWVCSIN-GTLQPGAVYSSISLCEFCRCLPQGPSPDAFVSCETQICN 723
 QY 267 V-----RPGQPV--YSLKKG-----KCKSK-- 286
 Db 724 THCPVGPEYQESGGCGCTGVQAVCTNTRSKSAHLFYPGETWSDAGNHCVTHQCEKHQ 783
 QY 287 -----TKSPPEV-----RF-----TYAGCS 302
 Db 784 GLVYVTKRACPLSLSLDEARMSKDCGCFPLPPPYONOSTCAVYHRSLLIIOOGCS 843
 QY 303 SVAKYKPYKC--GSCVVG-----RCCTPLOTRTVMKRFCEDEGEKSKVM 346
 Db 844 SSEPVRLATCRGCGSSSSSYSLGNTVEHRCCQCELRISLNAVTLHCTDGSRAFSIT 903
 QY 347 MIOCKC-NYNCPHNE 362
 Db 904 EVEECGCMGRCPAPGD 920

RESULT 14

A53767

mucin MUC5B, tracheobronchial - human

C:Species: Homo sapiens (man)

C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 20-Apr-2000

C:Accession: A53767

R.Meerzaman, D.; Charles, P.; Daskal, E.; Polymeropoulos, M.H.; Martin, B.M.; Rose, M.C.
 J. Biol. Chem. 269, 12932-12939, 1994
 A:Title: Cloning and analysis of cDNA encoding a major airway glycoprotein, human trachea
 A:Reference number: A53767; MUID:94230376; PMID:7513696
 A:Accession: A53767
 A:Status: Preliminary
 A:Molecule type: mRNA
 A:Residues: 1-1056 <ME>
 A:Cross-references: GB:006711; NID:9488072; PIDN:AA18431.1; PID:9488073
 A>Note: Authors translated the codon TAT for residue 1054 as Thr
 C:Genetics:
 A:Gene: GDB:MUC5B; MUC5
 A:Cross-references: GDB:128436; OMIM:600770
 A:Map position: 11p15.5-11p15.5
 C:Superfamily: von Willebrand factor type C repeat homology
 F:67-835/Domain: von Willebrand factor type C repeat homology <VWC>

Query Match 7.2%; Score 152; DB 2; Length 1056;
 Best Local Similarity 19.1%; Pred. No. 0.0035;
 Matches 107; Conservative 57; Mismatches 152; Indels 244; Gaps 30;

```

QY 4 STEFLAVATLHLRLALSTCPACHPLPAKAPGVLDGCG-----GC 52
Db 501 TTGSGTVGPTTVGTVGPTTPAPC---LPSPICHLILSKVEPCHTVIPPLTFYEGC 557
QY 53 -----CKVCAKQINEDC-----SKTOPCDH-----TKGLE----- 77
Db 558 VFDCCHMTDLDVYSSLEFLVARKLAS--HDICIDMRGRTYCAHHLPSRQGVPAIRPSNP 615
QY 78 --CNFGASSTALKGICRAQ--SEGRPCENSRIVON-----G 110
Db 616 SYC-XGNDASALGLAREGPIITGECFCEGMLFSTSAQVCVPFGCPRLGPHGEPYKVG 674
QY 111 ESFPNCKHQTG--IDGAVGCP-LCPQELSLPNLGCPNRLYKV-----SGOCCEEMV 162
Db 675 HTVGMDCQ-ECTCEAATWTLTCRPLCP---LPP-ACPLPGFVPVPAAPQAGCCCPQYS 728
QY 163 CDEDSINDSLDDQDGLGLDASEVEITFNNELIAGKSSSLKRLPVCTEERVLFPPLHA 222
Db 729 C-----ACNTSRCPAPVPCPSGA--RIPTY-----Q 753
QY 223 HGQKIVQVTSWSOC-----SKSCGT-----GISTRYTND 252
Db 754 EGACCPVQNCMTVCSINGTLXPGAVYSSSLCETCKRELPGPSDAFVSCETQICNT 813
QY 253 NPEGR-L-VKETRICEVRP-CGOPYSSILKG-----KCKSK----- 286
Db 814 HCPVRFETOEORRSVAVAPVCRSPVSPAPAPPTSTLASTWSDAGNHCVTHQCEKHODG 873
QY 287 -----TKSPRPV-----FTYAGC-----SS 303
Db 874 LVVYTTKACPLPCLSLDEAHMSKDGCCRCPLPPPPYONOSTCAVYHRSLLIIOQGSSS 933
QY 304 VKKTRPKYC-GSCVDG-----RCCTPLQTRTVKMRFRCEDEMFSSKNVM 347
Db 934 SEPRLAYCRNGCDSSSMYSLEGNTEVHRQCCQELRTSLRNTLIHCTDSSRAFSYTE 993
QY 348 IQSKC-NYNCPPHNEASF 366
Db 994 VEECGMGRCRCPAPAPSTR 1013

```

RESULT 15

S08167
 Balbiani ring 3 protein - midge (Chironomus tentans)
 C:Species: Chironomus tentans
 C:Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 21-Jul-2000
 C:Accession: S08167
 R:Paulsson, G.; Lendahl, U.; Gallit, J.; Ericsson, C.; Wieslander, L.
 J. Mol. Biol. 211, 331-349, 1990
 A:Title: The Balbiani ring 3 gene in Chironomus tentans has a diverged repetitive structure
 A:Reference number: S08167; MUID:90172404; PMID:1689777
 A:Accession: S08167

A:Status: not compared with conceptual translation
 A:Molecule type: DNA
 A:Residues: 1-1700 <PN>
 A:Cross-references: GB:X52263; NID:97057; PIDN:CAA36506.1; PID:97058
 C:Genetics:
 A:Gene: BR3
 A:Map position: 4
 C:Superfamily: unassigned Balbiani ring proteins

Query Match 7.2%; Score 151.5; DB 2; Length 1700;
 Best Local Similarity 22.7%; Pred. No. 0.0057;
 Matches 83; Conservative 41; Mismatches 143; Indels 99; Gaps 23;

```

QY 30 CHCLEAPKAPGVGLVNDGCG--C-CKVCAKQINEDSKTOPCDHTGLEGCNFGASSTA 86
Db 384 CIGPADAVCTAG-----KERGSECEKCIINREPEGAKP-----LVWN-----ENT 427
QY 87 LKICRAQSEGRPCENSRIVONGSEFQP-NCKHQCCTIDGAVG-----IPLCPQE 137
Db 428 CKVCPRADKQKSPGGCGS-----GKSFNKLTC--QCEBDQASAKGLKRWNAIDCKKCEQ 480
QY 138 LSLPNLGCPNRLV--KVSQCCEEWGDEDSIKDSLDDQDGLGLDASEVE-----LTR 190
Db 481 PGMPEEGCKQOTWISDKCKCECSPTITQAPQILD-----LNTCECKCPVNMIAQ 530
QY 191 NNELIAGKSSSLKRLPVCTEERVLFPPLHAHQKCIIVQTSWSOC-----SKSCGTGI 245
Db 531 KEKCKSPQWTDKSCICLCCSTTPATCEKQOTWCGAC-----QCICPGDKNKG--- 579
QY 246 STRVTNDNPECRIVKETRICEVRP-CGP-VYSSILKKGKSKTKKSPRYFTYAGCSS 303
Db 580 -NKKFEDKPSCECK-----CKNNPTCTSPQWADADDCCKCPKQKQKQ-----GGCDG 627
QY 304 VKKTRPKYCGSCVDGRCCCTPLQTRTVKMRFRCEDEMFSSKNVMIIQSKCNYNCPHNEA 363
Db 628 GQKWNDRVC-SC---GCPVP-----RPDCTNGQIYN-----INTCAGCGGIDKPCSP 670
QY 364 SFRLYS 369
Db 671 KQIITN 676

```

Search completed: August 5, 2003, 14:07:27
 Job time : 17.4592 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 5, 2003, 14:02:16; Search time 9.475 Seconds
(without alignments)
1881.068 Million cell updates/sec

Title: US-09-495-448a-2

Perfect score: 2103
Sequence: 1 MSSSTFRTLAVALTLHLTR.....PNEASFRLYSFLPNDIKFRD 379

Scoring table: BIOSUM62
Gapop 10.0, Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2103	100.0	379	1 CYR6_MOUSE	P18406 mus musculus
2	2079	98.9	379	1 CYR6_MOUSE	Q95722 rattus norv
3	1938	92.2	381	1 CYR6_HUMAN	O00622 homo sapien
4	1646.5	78.3	375	1 CE10_CHICK	P19336 gallus gall
5	960	45.6	347	1 CTGF_RAT	Q91828 rattus norv
6	957.5	45.5	348	1 CTGF_MOUSE	P29268 mus musculu
7	950	45.2	349	1 CTGF_HUMAN	P29279 homo sapien
8	937	44.6	349	1 CTGF_BOVIN	O18739 bos taurus
9	919	43.7	349	1 CTGF_PIG	O19113 sus scrofa
10	861	40.9	351	1 NOV_CHICK	P28666 gallus gall
11	856	40.7	353	1 NOV_COTTA	P42642 coturnix co
12	839.5	39.9	357	1 NOV_HUMAN	P48745 homo sapien
13	829.5	39.4	351	1 NOV_RAT	Q94245 rattus norv
14	827	39.3	354	1 NOV_MOUSE	O64239 mus musculu
15	824.5	39.2	343	1 NOV_XENLA	P51609 xenopus lae
16	526.5	25.0	251	1 CTGL_MOUSE	Q92064 mus musculu
17	524	24.9	250	1 CTGL_HUMAN	Q76076 homo sapien
18	492	23.4	250	1 CTGL_RAT	Q91828 rattus norv
19	158	7.5	447	1 NR1_MOUSE	Q92061 mus musculu
20	154	7.3	456	1 NR1_CHICK	Q90265 gallus gall
21	154	7.3	3110	1 LMA2_HUMAN	P24043 homo sapien
22	152.5	7.3	1233	1 LMA2_MOUSE	P88088 homo sapien
23	151.5	7.2	1700	1 BAR3_HUMAN	Q03376 chironomus
24	148	7.0	450	1 NR1_HUMAN	Q95440 homo sapien
25	146	6.9	837	1 WPC_RAT	P98089 rattus norv
26	145.5	6.9	2813	1 WPC_HUMAN	Q28295 canis fami
27	144.5	6.9	5703	1 LRP2_HUMAN	Q91828 rattus norv
28	144	6.8	4655	1 LRP2_MOUSE	Q91828 rattus norv
29	143	6.8	2482	1 WPC_PIG	Q28833 sus scrofa
30	141.5	6.7	1172	1 LMB3_HUMAN	Q93751 homo sapien
31	141	6.7	1587	1 LMB3_MOUSE	Q93751 homo sapien
32	141	6.7	5376	1 ZAN_MOUSE	O68799 mus musculu
33	140.5	6.7	2282	1 ZAN_RABBIT	P57999 oryctolagus

34	140	6.7	3106	1 LMA2_MOUSE	O60675 mus musculu
35	139.5	6.6	1178	1 TSP2_CHICK	P33440 gallus gall
36	139	6.6	473	1 FP2_MITGA	Q25464 mytilus gal
37	138	6.6	1964	1 MTC4_MOUSE	P31695 mus musculu
38	136.5	6.5	5179	1 MOC2_HUMAN	O02817 homo sapien
39	136	6.5	480	1 HRAL_MOUSE	O91118 mus musculu
40	135.5	6.4	305	1 ITP2_MOUSE	P47877 mus musculu
41	135.5	6.4	870	1 SMC2_HUMAN	O96966 homo sapien
42	135	6.4	3672	1 LML2_CAEL	Q21313 caenorhabdi
43	134	6.4	453	1 HRA3_HUMAN	P83110 homo sapien
44	134	6.4	677	1 SP87_DICDI	P54643 dictyosteli
45	133.5	6.3	424	1 SP60_DICDI	P15270 dictyosteli

ALIGNMENTS

RESULT 1	ID	CYR6_MOUSE	STANDARD:	PRT:	379 AA.
AC	P18406;				
DT	01-NOV-1990 (Rel. 16, Created)				
DT	01-NOV-1990 (Rel. 16, Last sequence update)				
DT	28-FEB-2003 (Rel. 41, Last annotation update)				
DE	CYR61 protein precursor (Cysteine-rich, angiogenic inducer, 61)				
DE	(Insulin-like growth factor-binding protein 10) (3GH1).				
GN	CYR61 OR IGFBP10.				
OS	Mus musculus (Mouse).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
OX	NCBI_TaxID=10090;				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=BALE/C; TISSUE=Fibroblast;				
RC	MEDLINE=90287146; PubMed=2355916;				
RA	O'Brien T.P., Yang G.P., Sanders L., Lau L.F.;				
RT	"Expression of cyr61, a growth factor-inducible immediate-early gene."				
RL	Mol. Cell. Biol. 10:3569-3577(1990).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=AJ; TISSUE=Embryonic fibroblast;				
RC	MEDLINE=91288203; PubMed=2062642;				
RA	Latinkic B.V., O'Brien T.P., Lau L.F.;				
RT	"Promoter function and structure of the growth factor-inducible immediate early gene cyr61."				
RL	Nucleic Acids Res. 19:3261-3267(1991).				
CC	-1- FUNCTION: MAY ACT AS ONE OF THE MANY GROWTH FACTOR-BINDING PROTEINS; PROMOTES PROLIFERATION, MIGRATION AND ADHESION.				
CC	-1- SUBCELLULAR LOCATION: Secreted.				
CC	-1- TISSUE SPECIFICITY: LOW IN KIDNEY, ADRENAL GLAND, TESTES, BRAIN, AND OVARY, MODERATE IN HEART, UTERUS, AND SKELETAL MUSCLE, HIGHEST IN LUNG.				
CC	-1- DEVELOPMENTAL STAGE: EXPRESSED FROM G(0)/G(1) THROUGH MID-G(1) IN NORMAL CELLS, AND AT A CONSTANT LEVEL IN RAPIDLY GROWING CELLS.				
CC	-1- INDUCTION: BY growth factors.				
CC	-1- SIMILARITY: Contains 1 IGF1R domain.				
CC	-1- SIMILARITY: Contains 1 WPC domain.				
CC	-1- SIMILARITY: Contains 1 TSP type-1 domain.				
CC	-1- SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain.				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@sib-sib.ch).				
CC	EMBL: M32490; AAA37512.1; -				
DR	EMBL: X56790; CAA40109.1; -				
DR	PIR: A35669; A35669.				
DR	MGI: 88613; Cyr61.				

DR GO: 0001569; P: patterning of blood vessels; IMP.
 DR InterPro: IPR006208; Cys_knot.
 DR InterPro: IPR006207; Cys_knot_C.
 DR InterPro: IPR000867; Insl_gro_fac_pr.
 DR InterPro: IPR000884; TSP1.
 DR InterPro: IPR001007; VWF_C.
 DR Pfam: PF00007; Cys_knot; 1.
 DR Pfam: PF00219; IGFBP; 1.
 DR Pfam: PF00090; tsp_1; 1.
 DR Pfam: PF00093; vwc; 1.
 DR SMART: SM00041; CT; 1.
 DR SMART: SM00121; IB; 1.
 DR SMART: SM00209; TSP1; 1.
 DR SMART: SM00214; VMC; 1.
 DR PROSITE: PS01185; CTCK_1; 1.
 DR PROSITE: PS01225; CTCK_2; 1.
 DR PROSITE: PS00222; IGF_BINDING; 1.
 DR PROSITE: PS50092; TSP1; 1.
 DR PROSITE: PS01208; VWF_C_1; 1.
 DR PROSITE: PS50184; VWF_C_2; 1.
 DR PROSITE: PS50184; VWF_C_2; 1.
 KM Growth factor binding; Signal.
 FT SIGNAL 1 24 POTENTIAL.
 FT CHAIN 25 379 CYR61 PROTEIN.
 FT DOMAIN 98 164 VWF.
 FT DOMAIN 226 271 TSP TYPE-1.
 FT DOMAIN 284 358 CTCK.
 FT DISULFID 321 321 BY SIMILARITY.
 FT DISULFID 301 335 BY SIMILARITY.
 FT DISULFID 312 351 BY SIMILARITY.
 FT DISULFID 315 353 BY SIMILARITY.
 FT DISULFID 320 357 BY SIMILARITY.
 SQ SEQUENCE 379 AA; 41709 MW; FA6B5014B56A8EB9 CRC64;

Query Match 100.0%; Score 2103; DB 1; Length 379;
 Best Local Similarity 100.0%; Pred. No. 5,6e-148;
 Matches 379; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSSTFRLAVAVTLHLTRLALSTCPAACHPEAPKCAPGVGLVVDGCGCCCKVCAKOL 60
 1 MSSSTFRLAVAVTLHLTRLALSTCPAACHPEAPKCAPGVGLVVDGCGCCCKVCAKOL 60

DB 1 MSSSTFRLAVAVTLHLTRLALSTCPAACHPEAPKCAPGVGLVVDGCGCCCKVCAKOL 60

QY 61 NECCSKTOPCDHRTKGLFCNFGASSSTALKGICRAQSEBGRPEVNSRYIQNESEFQPNCKHQ 120
 61 NECCSKTOPCDHRTKGLFCNFGASSSTALKGICRAQSEBGRPEVNSRYIQNESEFQPNCKHQ 120

DB 61 NECCSKTOPCDHRTKGLFCNFGASSSTALKGICRAQSEBGRPEVNSRYIQNESEFQPNCKHQ 120

QY 121 CTCIDGAVGCIPLCPQELSLPNLGCPRNRLVYKSGGCCCEWVDEDSIKSLDDODDLG 180
 121 CTCIDGAVGCIPLCPQELSLPNLGCPRNRLVYKSGGCCCEWVDEDSIKSLDDODDLG 180

DB 121 CTCIDGAVGCIPLCPQELSLPNLGCPRNRLVYKSGGCCCEWVDEDSIKSLDDODDLG 180

QY 181 LDASEVELFRNNELIAGKSSSLKRLPVFEGTEPRVLFNPLHAHGOKCIYQTSMSQCSKS 240
 181 LDASEVELFRNNELIAGKSSSLKRLPVFEGTEPRVLFNPLHAHGOKCIYQTSMSQCSKS 240

DB 181 LDASEVELFRNNELIAGKSSSLKRLPVFEGTEPRVLFNPLHAHGOKCIYQTSMSQCSKS 240

QY 241 CGGIGISRTYNDNPEGRVETRICVYRPGQPYVSSLLKGGKCSKTKSPREVRTYAG 300
 241 CGGIGISRTYNDNPEGRVETRICVYRPGQPYVSSLLKGGKCSKTKSPREVRTYAG 300

DB 241 CGGIGISRTYNDNPEGRVETRICVYRPGQPYVSSLLKGGKCSKTKSPREVRTYAG 300

QY 301 CSSVKKRRPKYCGSSCYDGRCTPLQTRTYVMRFRCEGEGEMSKNVMYIOSCKNYNCPHP 360
 301 CSSVKKRRPKYCGSSCYDGRCTPLQTRTYVMRFRCEGEGEMSKNVMYIOSCKNYNCPHP 360

DB 301 CSSVKKRRPKYCGSSCYDGRCTPLQTRTYVMRFRCEGEGEMSKNVMYIOSCKNYNCPHP 360

QY 361 NEASFRLYSLFNDIHKFRD 379
 361 NEASFRLYSLFNDIHKFRD 379

DB 361 NEASFRLYSLFNDIHKFRD 379

RESULT 2
 CYR6_RAT
 ID CYR6_RAT STANDARD; PRT; 379 AA.
 AC 09ES72;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE CYR61 protein precursor (Cysteine-rich, angiogenic inducer, 61)
 DE (Insulin-like growth factor-binding protein 10).
 GN CYR61 OR IGFBP10.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
 OX NCBI_Taxid=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lang;
 RX MEDLINE=20435857; PubMed=10852911;
 RA Albrecht C., von Der Kammer H., Mayhaus M., Klaudiny J., Schweizer M.,
 RA Nitsch R.M.;
 RT "Muscarinic acetylcholine receptors induce the expression of the
 RT immediate early growth regulatory gene CYR61.";
 RL J. Biol. Chem. 275:28929-28936(2000).
 CC -1- FUNCTION: MAY ACT AS ONE OF THE MANY GROWTH FACTOR-BINDING
 CC PROTEINS; PROMOTES PROLIFERATION, MIGRATION AND ADHESION (BY
 CC SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: Contains 1 IGFBP domain.
 CC -1- SIMILARITY: Contains 1 VWF domain.
 CC -1- SIMILARITY: Contains 1 TSP type-1 domain.
 CC -1- SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

DR EMBL: AF218558; AAC14964.1; -
 DR InterPro: IPR006208; Cys_knot.
 DR InterPro: IPR006207; Cys_knot_C.
 DR InterPro: IPR000867; Insl_gro_fac_pr.
 DR InterPro: IPR000884; TSP1.
 DR InterPro: IPR001007; VWF_C.
 DR Pfam: PF00007; Cys_knot; 1.
 DR Pfam: PF00029; IGFBP; 1.
 DR Pfam: PF00090; tsp_1; 1.
 DR Pfam: PF00093; vwc; 1.
 DR SMART: SM00041; CT; 1.
 DR SMART: SM00121; IB; 1.
 DR SMART: SM00209; TSP1; 1.
 DR SMART: SM00214; VMC; 1.
 DR PROSITE: PS01185; CTCK_1; 1.
 DR PROSITE: PS01225; CTCK_2; 1.
 DR PROSITE: PS00222; IGF_BINDING; 1.
 DR PROSITE: PS50092; TSP1; 1.
 DR PROSITE: PS01208; VWF_C_1; 1.
 DR PROSITE: PS50184; VWF_C_2; 1.
 KM Growth factor binding; Signal.
 FT SIGNAL 1 24 POTENTIAL.
 FT CHAIN 25 379 CYR61 PROTEIN.
 FT DOMAIN 98 164 VWF.
 FT DOMAIN 226 271 IGFBP.
 FT DOMAIN 284 358 TSP TYPE-1.
 FT DOMAIN 321 321 CTCK.
 FT DISULFID 301 335 BY SIMILARITY.
 FT DISULFID 312 351 BY SIMILARITY.
 FT DISULFID 315 353 BY SIMILARITY.
 FT DISULFID 320 357 BY SIMILARITY.
 SQ SEQUENCE 379 AA; 41687 MW; 62BF0BBA4C5AFDE9 CRC64;

Query Match 98.9%; Score 2079; DB 1; Length 379;
 Best Local Similarity 98.7%; Pred. No. 3.3e-146;
 Matches 374; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 MSSSTFRLAVAVTLHLTRLALSTCPAACHPEAPKCAPGVGLVVDGCGCCCKVCAKOL 60
 1 MSSSTFRLAVAVTLHLTRLALSTCPAACHPEAPKCAPGVGLVVDGCGCCCKVCAKOL 60

DB 1 MSSSTIKTILAVATLHLFRLALSTCPAACHCPLKAPGVLVNDGCGCCVCAKOL 60
 QY 61 NECCSKTQPCDHRKGLJECNCGASTALGICRAOSBERPCEYNSRITONESPQPNKHQ 120
 DB 61 NECCSKTQPCDHRKGLJECNCGASTALGICRAOSBERPCEYNSRITONESPQPNKHQ 120
 QY 121 CTCIDGAVGICPLCPQELSLPNIGCPNRLVYKSGCCCEWVCEDESKISLDDODLLG 180
 DB 121 CTCIDGAVGICPLCPQELSLPNIGCPNRLVYKSGCCCEWVCEDESKISLDDODLLG 180
 QY 181 LDASEVELLRNNELIAGKSSSLKRLPVGTEPRVLFNPLAHGQKCIQVOTSMSCSKS 240
 DB 181 FDASEVELLRNNELIAGKSSSLKRLPVGTEPRVLYNPLAHGQKCIQVOTSMSCSKS 240
 QY 241 CGAGISRTVYNDSECLVKEKTRICERPCGQPVYSLKKGKCKSTKSPPEVRYTAG 300
 DB 241 CGAGISRTVYNDSECLVKEKTRICERPCGQPVYSLKKGKCKSTKSPPEVRYTAG 300
 QY 301 CSSVKYRPRYKSCGVDRCTPLQRTVYKMRFCEDGEMFSKNVMMIOCKNCYNCPHP 360
 DB 301 CSSVKYRPRYKSCGVDRCTPLQRTVYKMRFCEDGEMFSKNVMMIOCKNCYNCPHP 360
 QY 361 NEASFRLYSLFNDIHKFRD 379
 DB 361 NEASFRLYSLFNDIHKFRD 379

RESULT 3
 CYR6_HUMAN STANDARD; PRT; 361 AA.
 ID CYR6_HUMAN 000622; 014934; 043775; 09BZL7;
 AC 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE CYR61 protein precursor (Cysteine-rich, angiogenic inducer, 61)
 DE (insulin-like growth factor-binding protein 10) (IGFBP10).
 GN CYR61 OR IGFBP10 OR GIG1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Albrecht C., von der Kammer H., Klaudiny J., Mayhaus M., Nitsch R.M.;
 RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA MEDLINE-97280750; PubMed-9135077;
 RA Jay P., Berge-Leiranc J.L., Marsollier C., Mejean C., Tavlaux S.,
 RA Berta P.;
 RL "the human growth factor-inducible immediate early gene, CYR61, maps
 to chromosome 1p.11-1753-1757(1997).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA MEDLINE-98197344; PubMed-9536281;
 RA Martinet C., Viegas-Pequignot E., Nguyen V.C., Perbal B.;
 RT "Chromosomal mapping and expression of the human cyr61 gene in tumour
 cells from the nervous system."
 RL Mol. Pathol. 50:310-316(1997).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RA Kolesnikova T.V., Lau L.F.;
 RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RA Bi A.B., Yu L.;
 RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A.
 RA Schuetze N., Lechner A., Groll C., Koehrl J., Jakob F.;
 RT "Regulation of hcyr61 by vitamin D, serum and cytokines in fetal human

RT osteoblasts.";
 RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
 RN [7]
 RP SEQUENCE FROM N.A.
 RA Leng E., Tai G., Estable M., Liu J., Chow C., Sadowski I.;
 RT "Organization and expression of the cyr61 gene in normal human
 RT fibroblasts.";
 RN Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
 RN [8]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lung, Placenta, and Skin;
 RX MEDLINE-22388257; PubMed-12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Bueltow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusik K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ugin T.B., Toshlyuk S., Carninci P., Prange C.,
 RA Raba S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Wotley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butlerfield R.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Mitra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -1- FUNCTION: MAY ACT AS ONE OF THE MANY GROWTH FACTOR-BINDING
 CC PROTEINS. PROMOTES PROLIFERATION, MIGRATION AND ADHESION (BY
 CC SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: Contains 1 IGFBP domain.
 CC -1- SIMILARITY: Contains 1 WFGC domain.
 CC -1- SIMILARITY: Contains 1 TSP type-1 domain.
 CC -1- SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain.
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 DR EMBL: Y12084; CAA72802.1; -;
 DR EMBL: U62015; AAB58319.1; -;
 DR EMBL: Y11307; CAA72167.1; -;
 DR EMBL: AF003594; AAB61240.1; -;
 DR EMBL: AF003185; AAB84227.1; -;
 DR EMBL: Z98053; CAB10848.1; -;
 DR EMBL: AF0307860; AAC59863.1; -;
 DR EMBL: BC001371; AAH01371.1; -;
 DR EMBL: BC009199; AAH09199.1; -;
 DR EMBL: BC016952; AAH16952.1; -;
 DR Genew: HGNC:2654; CYR61.
 DR MIM: 602369; -;
 DR GO: GO:0008283; P:cell proliferation; TAS.
 DR GO: GO:0007345; P:embryogenesis and morphogenesis; TAS.
 DR InterPro: IPR006208; Cys_knot.
 DR InterPro: IPR006207; Cys_knot_C.
 DR InterPro: IPR000867; Ins1_knot_C.
 DR InterPro: IPR000884; TSP1.
 DR InterPro: IPR001007; WFG_C.
 DR Pfam: PF00007; Cys_knot; 1.
 DR Pfam: PF00219; IGFBP; 1.
 DR Pfam: PF00090; TSP_1; 1.
 DR Pfam: PF00093; WFG; 1.
 DR SMART: SM00041; CT; 1.

DR SMART; SM00121; IB; 1.
 DR SMART; SM00209; TSP1; 1.
 DR SMART; SM00214; VWC; 1.
 DR PROSITE; PS01185; CTCK_1; 1.
 DR PROSITE; PS01225; CTCK_2; 1.
 DR PROSITE; PS00222; IGF_BINDING; 1.
 DR PROSITE; PS50092; TSP1; 1.
 DR PROSITE; PS01208; VWF_1; 1.
 DR PROSITE; PS50184; VWF_2; 1.
 KM Growth factor binding; Signal.
 FT SIGNAL 1 24
 FT CHAIN 25 381
 FT DOMAIN 26 97
 FT DOMAIN 98 164
 FT DOMAIN 228 273
 FT DOMAIN 286 360
 FT DOMAIN 286 323
 FT DISULFID 303 337
 FT DISULFID 314 353
 FT DISULFID 317 355
 FT DISULFID 322 359
 FT DISULFID 165 165
 FT CONFLICT 210 230
 FT CONFLICT 220 220
 FT CONFLICT 369 369
 SQ SEQUENCE 381 AA; 42026 MW; FC0BD39C078CA0B1 CRC64;

Query Match 92.2%; Score 1938; DB 1; Length 381;
 Best Local Similarity 91.4%; Pred. No. 8e-136;
 Matches 350; Conservative 9; Mismatches 18; Indels 6; Gaps 2;

QY 1 MSSSTFTLAVAVTLTLHLRLALSTCPACCHCPLEAPKCAPGVGLVDDCGCCCKVCAKOL 60
 DB 1 MSSIRARALATVTLTLHLRLALSTCPACCHCPLEAPKCAPGVGLVDDCGCCCKVCAKOL 60
 QY 61 NECCSTQPCDHTKGLGECNFGASSTALKGICRAQSGRCCEVNSRIYONGESFQPNCKKQ 120
 DB 61 NECCSTQPCDHTKGLGECNFGASSTALKGICRAQSGRCCEVNSRIYONGESFQPNCKKQ 120
 QY 121 CTCIDAVAGCIPCLPELSPNLGCPNPLVYKVGCCCEWVDEDSIKDSDDDDL-- 178
 DB 121 CTCIDAVAGCIPCLPELSPNLGCPNPLVYKVGCCCEWVDEDSIKDSDDDDL-- 178
 QY 121 CTCIDAVAGCIPCLPELSPNLGCPNPLVYKVGCCCEWVDEDSIKDSDDDDL-- 180
 DB 121 CTCIDAVAGCIPCLPELSPNLGCPNPLVYKVGCCCEWVDEDSIKDSDDDDL-- 180
 QY 179 --IGLDASEVELTRNNELIAGKSSILKRLPVFGTEPRVLFNPLAHAGKCIYQVTSWQ 236
 DB 181 KEGFADASEVELTRNNELIAGKSSILKRLPVFGTEPRVLFNPLAHAGKCIYQVTSWQ 238
 QY 237 CSKSCGTGISTRTNPNPCRLVKEIRICEVPRCGPVYSSLLKKGKCKSKTKSPPEVRF 296
 DB 239 CSKTCGTGISTRTNPNPCRLVKEIRICEVPRCGPVYSSLLKKGKCKSKTKSPPEVRF 298
 QY 297 TYAGCSVKKRYKRYKYGSCVDCGCTPLQTRPKMFRCEDEGEMFSKNYMMIOSCKCNYN 356
 DB 299 TYAGCLSTVKKRYKRYKYGSCVDCGCTPLQTRPKMFRCEDEGEMFSKNYMMIOSCKCNYN 358
 QY 357 CPHPEASFRLYSLFNDHKFRD 379
 DB 359 CPHANEAPFRYLFNDHKFRD 381

RESULT 4
 CE10_CHICK STANDARD; PRT; 375 AA.
 AC P19336;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE CEF-10 protein precursor.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OK NCBI_TaxID=9031;

RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89145206; PubMed=2537491;
 RA Simmons D.L., Levy D.B., Yannoni Y., Erikson R.L.;
 RT Identification of a phorbol ester-repressible v-src-inducible gene.;
 RL Proc. Natl. Acad. Sci. U.S.A. 86:1178-1182(1989).
 CC -1- FUNCTION: PROBABLE SECRETED REGULATORY PROTEIN.
 CC -1- SUBCELLULAR LOCATION: Secreted (Probable).
 CC -1- INDUCTION: By v-src.
 CC -1- SIMILARITY: Contains 1 IGF1R domain.
 CC -1- SIMILARITY: Contains 1 VWF domain.
 CC -1- SIMILARITY: Contains 1 TSP type-1 domain.
 CC -1- SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-slb.ch/announce/>
 CC or send an email to license@slb.ch).
 CC -----
 CC EMBL: J04496; AAA48661.1; -
 CC PIR: A41428; A41428.
 CC InterPro: IPR006208; Cys_knot.
 CC InterPro: IPR006207; Cys_knot.C.
 CC InterPro: IPR000867; Ins1_gro_fac-pr.
 CC InterPro: IPR000884; TSP1.
 CC InterPro: IPR001007; VWF.C.
 CC Pfam: PF00007; Cys_knot; 1.
 CC Pfam: PF00219; IGF1R; 1.
 CC Pfam: PF00090; TSP_1; 1.
 CC Pfam: PF00093; VWC; 1.
 DR SMART; SM00041; CT; 1.
 DR SMART; SM00121; IB; 1.
 DR SMART; SM00209; TSP1; 1.
 DR SMART; SM00214; VWC; 1.
 DR PROSITE; PS01185; CTCK_1; 1.
 DR PROSITE; PS01225; CTCK_2; 1.
 DR PROSITE; PS00222; IGF_BINDING; 1.
 DR PROSITE; PS50092; TSP1; 1.
 DR PROSITE; PS01208; VWF_1; 1.
 DR PROSITE; PS50184; VWF_2; 1.
 KW Growth factor binding; Signal.
 FT SIGNAL 1 22
 FT CHAIN 23 375
 FT DOMAIN 98 164
 FT DOMAIN 223 268
 FT DOMAIN 281 355
 FT DISULFID 281 318
 FT DISULFID 298 332
 FT DISULFID 309 348
 FT DISULFID 312 350
 FT DISULFID 317 354
 SQ SEQUENCE 375 AA; 40651 MW; 95F28533BE35D5AE CRC64;

Query Match 78.3%; Score 1646.5; DB 1; Length 375;
 Best Local Similarity 80.1%; Pred. No. 2.3e-114;
 Matches 302; Conservative 22; Mismatches 36; Indels 17; Gaps 7;

QY 10 AAVATLHLTRLALSTCPACCHCPLEAPKCAPGVGLVDDCGCCCKVCAKOLNECCSKQ 68
 DB 9 ALAAALCLARLAGSPAVCOCPAAAPQCAPGVGLVDDCGCCCKVCAKOLNECCSKQ 68
 QY 69 PCDHTKGLGECNFGASSTALKGICRAQSGRCCEVNSRIYONGESFQPNCKHCTCIDGAV 128
 DB 69 PCDHTKGLGECNFGASSTALKGICRAQSGRCCEVNSRIYONGESFQPNCKHCTCIDGAV 128
 QY 129 GCIPCLPELSPNLGCPNPLVYKVGCCCEWVDEDSIKDSDDDDL--IGLDAS 184
 DB 129 GCIPCLPELSPNLGCPNPLVYKVGCCCEWVDEDS--KALDELBEFFSKREGGLAS 186
 QY 185 EVELTRNNELIAGKSSILKRLPVFGTEPRVLFNPLAHAGKCIYQVTSWQSKSCG 242

```

Db 187 EGEELRNNEILAIYK-G-GLKMLPVFGSEPOSRAPENP-----KCIYQTTSMSCSKTCG 239
QY 243 TGISTRTYNDNPECRIVKTRICEVRPCGQPVYSSLLKGGKCKSKTSKSPREVRTYAGCS 302
Db 240 TGISTRTYNDNPECRIVKTRICEVRPCGQPVYSSLLKGGKCKSKTSKSPREVRTYAGCS 299
QY 303 SVKRYRKYCGSCVDGRCCTPLQRTYKMRFCDEGEMFSKNVMIOSCKNVCNCPHNE 362
Db 300 SVKRYRKYCGSCVDGRCCTPLQRTYKMRFCDEGEMFSKNVMIOSCKNVCNCPHNE 359
QY 363 ASFRLYSLFNDIHKFRD 379
Db 360 A-YFRLVNDIHKFRD 375

RESULT 5
CTGF_RAT STANDARD: PRT: 347 AA.
ID CTGF_RAT STANDARD: PRT: 347 AA.
AC O9R1E9: 09WVSL:
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Connective tissue growth factor precursor.
GN CTGF.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20145935; PubMed=10679821;
RA Xu J., Smock S.L., Safadi F.F., Rosenzweig A.B., Odgren P.R.,
RT Marks S.C. Jr., Owen T.A., Popoff S.N.:
RT "Cloning the full-length cDNA for rat connective tissue growth factor:
RT implications for skeletal development.";
RL J. Cell. Biochem. 77:103-115(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Teruka K., Tamatani T.:
RT "Rattus norvegicus connective tissue growth factor.";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: MAJOR CONNECTIVE TISSUE MITOATRACTANT SECRETED BY
CC VASCULAR ENDOTHELIAL CELLS. PROMOTES PROLIFERATION AND
CC DIFFERENTIATION OF CHONDROCYTES (BY SIMILARITY). MEDIATES CELL
CC ADHESION AND ENHANCES FIBROBLAST GROWTH FACTOR-INDUCED DNA
CC SYNTHESIS (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: Contains 1 IGFBP domain.
CC -1- SIMILARITY: Contains 1 WPC domain.
CC -1- SIMILARITY: Contains 1 TSP type-1 domain.
CC -1- SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF120275; BAB39132.1;
DR EMBL: AB023068; BAB82125.1;
DR InterPro: IPR006208; Cys_knot.
DR InterPro: IPR006207; Cys_knot_C.
DR InterPro: IPR000867; Ins1_gro_fac_pr.
DR InterPro: IPR000884; TSP1.
DR InterPro: IPR001007; WVF.C.
DR Pfam: PF00007; Cys_knot; 1.
DR Pfam: PF00219; IGFBP; 1.
DR Pfam: PF00090; TSP_1; 1.
DR Pfam: PF00093; WVF; 1.
DR SMART: SM00041; CT; 1.

```

```

DR SMART: SM0021; IB; 1.
DR SMART: SM00209; TSP1; 1.
DR SMART: SM00214; WVF; 1.
DR PROSITE: PS01185; CTCK_1; 1.
DR PROSITE: PS01225; CTCK_2; 1.
DR PROSITE: PS00222; IGFBP; 1.
DR PROSITE: PS00092; TSP1; 1.
DR PROSITE: PS01208; WVF; 1.
DR PROSITE: PS0184; WVF; 2.
DR Cell adhesion; DNA synthesis; Extracellular matrix; Signal.
KM SIGNAL
FT CHAIN 1 24
FT CHAIN 25 347
FT CHAIN 31 98
FT CHAIN 99 165
FT CHAIN 196 241
FT CHAIN 254 328
FT CHAIN 254 291
FT CHAIN 271 305
FT CHAIN 282 321
FT CHAIN 285 323
FT CHAIN 290 327
FT CHAIN 35 35
FT CHAIN 94 94
FT CHAIN 94 94
FT CHAIN 347 AA; 37756 MW; CFEI1A197687B16 CRC64;
SQ SEQUENCE

Query Match. 45.6%; Score 960; DB 1; Length 347;
Best Local Similarity 46.4%; Pred. No. 7.5e-64;
Matches 176; Conservative 61; Mismatches 104; Indels 38; Gaps 7;

QY 1 MSSSTFRLAVAVTLHLRLAL-STCPACCHPLE-APKCAPGVGLVRCGCGCKVCAR 58
Db 1 MLASVAPVSLVALVLLCTRPATGQDCSAQCAEAAPRCBPAGVSLVLDGCGCCRVAC 60
QY 59 QLNDCSCKTPQPCDHITKLECFGASSSTALKICCAQSGRCCEFNRYONGESFQPNCK 118
Db 61 QLGELCTERPCDCHKGLFCDFGSPANKRIGVCTAK-DGACVGVGSYRSGESFQSSCK 119
QY 119 HOCCTIGAVGCIPLCPQELSLPMLGCPNPLVAVSGCCCEWCEDESIDDDODL 178
Db 120 YQCTCLDGAAGCVPLCSMDVRLPSPDCFFPRVRLPGKCEWYCDER----- 167
QY 179 LGLDASEVELLRNNEILAIYKSSILKRL-PVFGTEPRVLPNLAHAGKCIYQTTSMSC 237
Db 168 -----KDRIVVGPAALAYRLIEDFFGDPIMM-----RANLCVQTTSMSC 207
QY 238 SKSCGTGISTRTYNDNPECRIVKTRICEVRPCGQPVYSSLLKGGKCKSKTSKSPREVRT 297
Db 208 SKTCGMGISTRTYNDNPECRIVKTRICEVRPCGQPVYSSLLKGGKCKSKTSKSPREVRT 267
QY 298 YAGGSVKKYRKYCGSCVDGRCCTPLQRTYKMRFCDEGEMFSKNVMIOSCKNVCN 357
Db 268 LSGCTSVKTYRAKCGCTGRCCTPHRTTLVPEFKCPGELMKKMMFLKTCACHYNC 327
QY 358 PHPNEASFRLY--SLFNDI 374
Db 328 PGDMDFRESILYRKMGDM 346

RESULT 6
CTGF_MOUSE STANDARD: PRT: 348 AA.
ID CTGF_MOUSE STANDARD: PRT: 348 AA.
AC P29268: 0922U0:
DT 01-DEC-1992 (Rel. 24, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Connective tissue growth factor precursor (FISP-12 protein)
DE (Hypertrophic chondrocyte-specific protein 24).
GN CTGF OR FISP12 OR FISP-12 OR HCS24.
OS Mus musculus (mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]

```

RP SEQUENCE FROM N.A.
 RX MEDLINE-91363290; PubMed-1888698;
 RA Ryseck R.-P., Macdonald-Bravo H., Mattei M.-G., Bravo R.;
 RT "Structure, mapping, and expression of fisp-12, a growth factor-
 RT inducible gene encoding a secreted cysteine-rich protein.";
 RL Cell Growth Differ. 2:225-233(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-91229699; PubMed-2029337;
 RA Brunner A., Chinn J., Neubauer M.G., Purchio A.F.;
 RT "Identification of a gene family regulated by transforming growth
 RT factor-beta.";
 RL DNA Cell Biol. 10:293-300(1991).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-22388257; PubMed-12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Dlatshenko L., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Bask S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gumarane P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulys S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shcherchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [4]
 RP FUNCTION AND SUBCELLULAR LOCATION.
 RX MEDLINE-97327410; PubMed-9184077;
 RA Kireeva M.L., Latinkic B.V., Kolesnikova T.V., Chen C.C., Yang G.P.,
 RA Adler A.S., Lau L.F.;
 RT "Cyr61 and fisp2 are both ECM-associated signaling molecules:
 RT activities, metabolism, and localization during development.";
 RL Exp. Cell Res. 233:63-77(1997).
 RN [5]
 RP FUNCTION.
 RX MEDLINE-99182484; PubMed-10082563;
 RA Babic A.M., Chen C.C., Lau L.F.;
 RT "Fisp2/mouse connective tissue growth factor mediates endothelial
 RT cell adhesion and migration through integrin alphavbeta3, promotes
 RT endothelial cell survival, and induces angiogenesis in vivo.";
 RL Mol. Cell. Biol. 19:2958-2966(1999).
 CC -1- FUNCTION: MAJOR CONNECTIVE TISSUE MITOATRACTANT SECRETED BY
 CC VASCULAR ENDOTHELIAL CELLS. PROMOTES PROLIFERATION AND
 CC DIFFERENTIATION OF CHONDROCYTES (BY SIMILARITY). MEDIATES CELL
 CC ADHESION AND ENHANCES FIBROBLAST GROWTH FACTOR-INDUCED DNA
 CC SYNTHESIS (BY SIMILARITY).
 CC -1- SUBUNIT: Monomer (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: FOUND IN THE EXTRACELLULAR MATRIX AND AS A
 CC SOLUBLE FORM.
 CC -1- TISSUE SPECIFICITY: TESTIS, SPLEEN, KIDNEY, LUNG, HEART, AND BRAIN
 CC (LOWEST LEVEL IN TESTIS AND HIGHEST IN LUNG).
 CC -1- INDUCTION: By growth factors.
 CC -1- SIMILARITY: Contains 1 IGFBP domain.
 CC -1- SIMILARITY: Contains 1 WFC domain.
 CC -1- SIMILARITY: Contains 1 TSP type-1 domain.
 CC -1- SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M70641; AAA37627.1; -;
 CC DR EMBL: M70642; AAA37628.1; -;
 CC DR EMBL: M80263; AAA73135.1; -;
 CC DR EMBL: BC006783; AA06783.1; -;
 CC DR PIR: A40578; AA0578.
 CC DR MGI: 95537; Ctgf.
 CC DR GO: 0005578; C:extracellular matrix; IDA.
 CC DR GO: 0008201; F:heparin binding activity; IDA.
 CC DR GO: 0005178; F:integrin binding activity; IDA.
 CC DR GO: 0001525; P:angiogenesis; IDA.
 CC DR GO: 0001647; P:cell migration; IDA.
 CC DR GO: 0007160; P:cell-matrix adhesion; IDA.
 CC DR GO: 0008543; P:FGF receptor signaling pathway; IDA.
 CC DR GO: 0007229; P:integrin-mediated signaling pathway; IDA.
 CC DR GO: 0001503; P:proliferation; IMP.
 CC DR InterPro: IPR006208; Cys_knot.
 CC DR InterPro: IPR006207; Cys_knot_C.
 CC DR InterPro: IPR000867; Insl_gro_fac-pr.
 CC DR InterPro: IPR000884; TSP1.
 CC DR InterPro: IPR001007; WFC_C.
 CC DR Pfam: PF00007; Cys_knot; 1.
 CC DR Pfam: PF00219; IGFBP; 1.
 CC DR Pfam: PF00090; tsp.1; 1.
 CC DR Pfam: PF00093; wfc; 1.
 CC DR SMART: SM00041; CT; 1.
 CC DR SMART: SM00121; IB; 1.
 CC DR SMART: SM00209; TSP1; 1.
 CC DR SMART: SM00214; WVC; 1.
 CC DR PROSITE: PS01185; CTCK_1; 1.
 CC DR PROSITE: PS01225; CTCK_2; 1.
 CC DR PROSITE: PS00222; IGF_BINDING; 1.
 CC DR PROSITE: PS00082; TSP1; 1.
 CC DR PROSITE: PS01208; WFC_1; 1.
 CC DR PROSITE: PS0184; WFC_2; 1.
 CC DR Cell adhesion; DNA synthesis; Extracellular matrix; Signal.
 CC FT CHAIN; 1; 25; POTENTIAL.
 CC FT SIGNAL; 26; 348; CONNECTIVE TISSUE GROWTH FACTOR.
 CC FT DOMAIN; 32; 99; IGFBP.
 CC FT DOMAIN; 100; 166; WFC.
 CC FT DOMAIN; 197; 242; TSP TYPE-1.
 CC FT DOMAIN; 255; 329; CTCK.
 CC FT DISULFID; 255; 292; BY SIMILARITY.
 CC FT DISULFID; 272; 306; BY SIMILARITY.
 CC FT DISULFID; 283; 322; BY SIMILARITY.
 CC FT DISULFID; 286; 324; BY SIMILARITY.
 CC FT DISULFID; 291; 328; BY SIMILARITY.
 CC FT CONFLICT; 161; 161; E -> K (IN REF. 1).
 CC SQ SEQUENCE 348 AA; 37794 MW; 4D7B6D9089174049 CRC64;
 CC -----
 CC Query Match 45.5%; Score 957.5; DB 1; Length 348;
 CC Best Local Similarity 46.8%; Pred. No. 1; le-63;
 CC Matches 178; Conservative 60; Mismatches 103; Indels 39; Gaps 8;
 CC -----
 CC 1 MESSRTTAAVAVVILHL-TRLAL-STCPAACHCLE-APKCAPGGLVRODCGCKCYA 57
 CC 1 MASVAGPISLALVLLALCTRPATGDCSAQCQAAEAAPHCAGVSLVLDGCGCRCA 60
 CC 58 KQINEDCKTOPCDHTKGLGECNFGASSTALKICICRAOSGRCEYNSRYONGSEFOPNC 117
 CC 61 KQLGELCTRPDCDPKHLGFOFGSPANKKIGVCTAK-DGACVCGSVYRGESEFQSSC 119
 CC 118 KQCTCIGDAGVCCIPCLPQELSLPNLGCNPPLVVSQGCCEWVCDEDSIKDSLDDDD 177
 CC 120 KYQCTCLGAGVCPVCSMDVRLPSPDCPFPRRYVLPKCCBEWVCDEP----- 168
 CC 178 LIGLDASVEVLELRNNELIAGKSSLKLRL-PVEGTEPPVLPNPLAHOKCIYOTTSQ 236
 CC 169 -----KRTAVGPALAAVRYEDTGGDPYTM-----RANCLVOTTEMSA 207
 CC 237 CSKSGGTGISTRTVNDNPECRLVKTETRICEVAPCGQPYVSSILKKGKSKTKKSPPEVRF 296


```

Db 208 CSKTCGMSISTRYNDWTFCELEKOSRLCMWRPCPCADLEENIKKGGKCC:RTPKIAKPYKE 267
Oy 297 TYACSSVKKRYRPRYCCSCVDGRCTPLQRTYVAMRRRCDEGEFNSNNYAMIGCKKNYN 356
Db 268 ELSGCTSKYKRYRAFCVCTDGRCTPHRTTLPVEKRCPDGELMKRNKMEIKYCACHYN 327
Oy 357 CPHNEASFRLY--SLFNFI 374
Db 328 CPGDNDIFESLYIKRMKMGDM 347

RESULT 7
CTGF_HUMAN
ID CTGF_HUMAN STANDARD; PRT: 349 AA.
AC P29279; Q960X2;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Connective tissue growth factor precursor (Hypertrophic chondrocyte-
DE specific protein 24).
CN CTGF OR HCS24.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo
OX NCBI_Taxid:9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Umbilical vein endothelial cells;
RX MEDLINE=91373462; PubMed=1654338;
RA Bradham D.M., Igarashi A., Potter R.L., Grotendorst G.R.;
RT "Connective tissue growth factor: a cycline-rich mitogen secreted by
RT human vascular endothelial cells is related to the SRC-induced
RT immediate early gene product CEF-10."
RL J. Cell Biol. 114:1285-1294(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE-Umbilical vein endothelial cells;
RX MEDLINE=93187114; PubMed=1293144;
RA Igarashi A., Bradham D.M., Okochi H., Grotendorst G.R.;
RT "Connective tissue growth factor."
RL J. Dermatol. 19:642-643(1992).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=97207446; PubMed=9054739;
RA Oemar B.S., Werner A., Garnier J.M., Do D.D., Godoy N., Nuuck M.,
RA Marz W., Rupp J., Pech M., Luescher T.F.;
RT "Human connective tissue growth factor is expressed in advanced
RT atherosclerotic lesions."
RL Circulation 95:831-839(1997).
RN [4]
RP SEQUENCE FROM N.A.
RA COoley V.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RN [5]
RP FUNCTION.
RC TISSUE-Chondrocytes;
RX MEDLINE=20080284; PubMed=10614647;
RA Nakanishi T., Nishida T., Shimo T., Kobayashi K., Kubo T.,
RA Tamatani T., Tezuka K., Takigawa M.;
RT "Effects of CTGF/HCS24, a product of a hypertrophic chondrocyte-
RT specific gene, on the proliferation and differentiation of
RT chondrocytes in culture."
RL Endocrinology 141:264-273(2000).
RN [6]
RP FUNCTION: MAJOR CONNECTIVE TISSUE MITOATTRACTANT SECRETED BY
RN [7]
RP VASCULAR ENDOTHELIAL CELLS. PROMOTES PROLIFERATION AND
RN [8]
RP DIFFERENTIATION OF CHONDROCYTES. MEDIATES CELL ADHESION AND
RN [9]
RP ENHANCES FIBROBLAST GROWTH FACTOR-INDUCED DNA SYNTHESIS.
RN [10]
RP SUBUNIT: Monomer.
RN [11]
RP SUBCELLULAR LOCATION: FOUND IN THE EXTRACELLULAR MATRIX AND AS A
RN [12]
RP SOLUBLE FORM (BY SIMILARITY).
RN [13]
RP ALTERNATIVE PRODUCTS:
RN [14]
RP Event-Alternative splicing; Named isoforms=2;

```

```

CC Name=Long;
CC IsoId=P29279-1; Sequence=Displayed;
CC Name=Short;
CC IsoId=P29279-2; Sequence=VSP_002460;
CC Note-No experimental confirmation available;
CC -1- SIMILARITY: Contains 1 IGFBP domain.
CC -1- SIMILARITY: Contains 1 WFEC domain.
CC -1- SIMILARITY: Contains 1 TSP type-1 domain.
CC -1- SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M92934; AAA91279.1;
DR EMBL: X78947; CAA5544.1;
DR EMBL: A1354866; CAC44023.1;
DR PIR: A40551; A40551.
DR Genew: HGNC:2500; CTGF.
DR MIM: 121009;
DR GO: GO:0005578; C:extracellular matrix; TAS.
DR GO: GO:0005886; C:plasma membrane; TAS.
DR GO: GO:0005520; F:insulin-like growth factor binding activity; TAS.
DR GO: GO:0008151; P:cell growth and/or maintenance; TAS.
DR GO: GO:0008544; P:epidermal differentiation; TAS.
DR GO: GO:0009611; P:response to wounding; TAS.
DR InterPro: IPR006208; Cys_knot.
DR InterPro: IPR006207; Cys_knot_C.
DR InterPro: IPR000867; Insl_gro_fac_pr.
DR InterPro: IPR000884; TSP1.
DR InterPro: IPR001007; WFEC.
DR Pfam: PF00007; Cys_knot_1.
DR Pfam: PF00219; IGFBP_1.
DR Pfam: PF00090; tsp_1; 1.
DR Pfam: PF00093; WVC_1.
DR SMART: SM00041; CT; 1.
DR SMART: SM00121; IB; 1.
DR SMART: SM00209; TSP1; 1.
DR SMART: SM00214; WVC; 1.
DR PROSITE: PS01185; CTCK_1; 1.
DR PROSITE: PS01225; CTCK_2; 1.
DR PROSITE: PS00292; IGF_BINDING; 1.
DR PROSITE: PS00092; TSP1; 1.
DR PROSITE: PS01208; WFEC_1; 1.
DR PROSITE: PS0184; WFEC_2; 1.
KW Cell adhesion; DNA synthesis; Extracellular matrix; Signal;
KW Alternative splicing.
FT SIGNAL 1 26
FT CHAIN 27 349
FT DOMAIN 33 100
FT DOMAIN 101 167
FT DOMAIN 198 243
FT DOMAIN 256 330
FT DISULFID 256 293
FT DISULFID 273 307
FT DISULFID 284 323
FT DISULFID 287 325
FT DISULFID 292 329
FT CARBOHYD 28 28
FT CARBOHYD 225 225
FT VARSPPLIC 172 198
FT CONFLICT 83 83
FT FT 349 AA; 38069 MM; DECF8470B357EA95 CRC64;
SQ SEQUENCE 349 AA; 38069 MM; DECF8470B357EA95 CRC64;
Query Match 45.28; Score 950; DB 1; Length 349;
Best Local Similarity 45.78; Pred. No. 4; le-63;
Matches 176; Conservative 64; Mismatches 97; Indels 48; Gaps 9;

```

```

QY 1 MSSSTFRTLAVALTLHLTRIALSTCPAA-----CHCPLE-APKCAPGVGLVDRGCCG 52
DB 1 MTAASMPRVAVFAVYL-----LALCSRPVAGONSGPCRCDEDPAPRCAGVSLVLDGCCG 56
QY 53 CRYCAKOLNEDSCSTOPOCDHTKGLCECNFGASSALKGICCAOSEGRCEYNSRIYONGES 112
DB 57 CRYCAKOLNEDSCSTOPOCDHTKGLCECNFGASSALKGICCAOSEGRCEYNSRIYONGES 115
QY 113 FQPNCKHOCTCIDGAVGICPLCPQELSLPNLGCNPNRLVYSGOCCEWYCDSDISL 172
DB 116 FQSSCKYQCTCIDGAVGICPLCPQELSLPNLGCNPNRLVYSGOCCEWYCDSDISL 170
QY 173 DDQDGLGLDASEVELETRNNELIAGKSSILKRL-PVETGEPRVLEPNLHAHQKCIYOT 231
DB 171 DDQDGLGLDASEVELETRNNELIAGKSSILKRL-PVETGEPRVLEPNLHAHQKCIYOT 203
QY 232 TSMSCSGSGTGISTRTVNDNPECRVYKTRICEVRPGQPYVSSLKKGKSKTKKSP 291
DB 204 TSMSCSGSGTGISTRTVNDNPECRVYKTRICEVRPGQPYVSSLKKGKSKTKKSP 263
QY 292 EPRFTYAGSSVKKYRPKYCGSCVDRCCTPLQTRTKKRFCEDEGEMFSKNVMIOSG 351
DB 264 KPIKFLSGCTSKMKTFRANKCGVCTDGRCTPHRTTLVPEFRCPDEGVKKMMFRTIC 323
QY 352 KCNYNCPHNEASFRIL--SLFNDI 374
DB 324 ACHYNCPGNDIFESLYRKMGM 348

```

RESULT 8

CTGF_BOVIN STANDARD; PRT; 349 AA.

```

AC 018739; 09GL71;
DT 15-JUL-1998 (Rel. 36, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DT Connective tissue growth factor precursor.
GN CTGF.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN 11
RP SEQUENCE FROM N.A.
RC TISSUE=Arteria;
RA Lillensiek B., Lin Z., Fotsis T., Schimanski M., Bierhaus A.,
RA Kanitz M., Kauffmann G., Schweigerer L., Ziegler R., Nawroth P.P.;
RL Submitted (Aug-1997) to the EMBL/GenBank/DBJ databases.
RN 12
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Mathias M., Schwitters C., Hove M., Rupp S., Erondu N.E.;
RT "Bovine connective tissue growth factor, organization of the
RT chromosomal gene and demonstration of promoter activity.";
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: MAJOR CONNECTIVE TISSUE MITOATRACTANT SECRETED BY
CC VASCULAR ENDOTHELIAL CELLS. PROMOTES PROLIFERATION AND
CC DIFFERENTIATION OF CHONDROCYTES (BY SIMILARITY). MEDIATES CELL
CC ADHESION AND ENHANCES FIBROBLAST GROWTH FACTOR-INDUCED DNA
CC SYNTHESIS (BY SIMILARITY).
CC -1- SUBUNIT: Monomer (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: FOUND IN THE EXTRACELLULAR MATRIX AND AS A
CC SOLUBLE FORM (BY SIMILARITY).
CC -1- SIMILARITY: Contains 1 IGFBP domain.
CC -1- SIMILARITY: Contains 1 WFC domain.
CC -1- SIMILARITY: Contains 1 TSP type-1 domain.
CC -1- SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

```

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib.ch).
CC -----
DB 1 MSSSTFRTLAVALTLHLTRIALST-----CPAACHPPL-EAPRCAPGVGLVDRGCCG 52
DB 1 MSATGIGPVRCAFEVL-----LALCSRPVAGONSGPCRCDEDPAPRCAGVSLVLDGCCG 56
QY 53 CRYCAKOLNEDSCSTOPOCDHTKGLCECNFGASSALKGICCAOSEGRCEYNSRIYONGES 112
DB 57 CRYCAKOLNEDSCSTOPOCDHTKGLCECNFGASSALKGICCAOSEGRCEYNSRIYONGES 115
QY 113 FQPNCKHOCTCIDGAVGICPLCPQELSLPNLGCNPNRLVYSGOCCEWYCDSDISL 172
DB 116 FQSSCKYQCTCIDGAVGICPLCPQELSLPNLGCNPNRLVYSGOCCEWYCDSDISL 169
QY 173 DDQDGLGLDASEVELETRNNELIAGKSSILKRL-PVETGEPRVLEPNLHAHQKCIYOT 231
DB 170 DDQDGLGLDASEVELETRNNELIAGKSSILKRL-PVETGEPRVLEPNLHAHQKCIYOT 203
QY 232 TSMSCSGSGTGISTRTVNDNPECRVYKTRICEVRPGQPYVSSLKKGKSKTKKSP 291
DB 204 TSMSCSGSGTGISTRTVNDNPECRVYKTRICEVRPGQPYVSSLKKGKSKTKKSP 263
QY 292 EPRFTYAGSSVKKYRPKYCGSCVDRCCTPLQTRTKKRFCEDEGEMFSKNVMIOSG 351

```

Query Match 44.6%; Score 937; DB 1; Length 349;
Best Local Similarity 44.7%; Pred. No. 3.7e-62;
Matches 172; Conservative 65; Mismatches 100; Indels 48; Gaps 8;

```

QY 1 MSSSTFRTLAVALTLHLTRIALST-----CPAACHPPL-EAPRCAPGVGLVDRGCCG 52
DB 1 MSATGIGPVRCAFEVL-----LALCSRPVAGONSGPCRCDEDPAPRCAGVSLVLDGCCG 56
QY 53 CRYCAKOLNEDSCSTOPOCDHTKGLCECNFGASSALKGICCAOSEGRCEYNSRIYONGES 112
DB 57 CRYCAKOLNEDSCSTOPOCDHTKGLCECNFGASSALKGICCAOSEGRCEYNSRIYONGES 115
QY 113 FQPNCKHOCTCIDGAVGICPLCPQELSLPNLGCNPNRLVYSGOCCEWYCDSDISL 172
DB 116 FQSSCKYQCTCIDGAVGICPLCPQELSLPNLGCNPNRLVYSGOCCEWYCDSDISL 169
QY 173 DDQDGLGLDASEVELETRNNELIAGKSSILKRL-PVETGEPRVLEPNLHAHQKCIYOT 231
DB 170 DDQDGLGLDASEVELETRNNELIAGKSSILKRL-PVETGEPRVLEPNLHAHQKCIYOT 203
QY 232 TSMSCSGSGTGISTRTVNDNPECRVYKTRICEVRPGQPYVSSLKKGKSKTKKSP 291
DB 204 TSMSCSGSGTGISTRTVNDNPECRVYKTRICEVRPGQPYVSSLKKGKSKTKKSP 263
QY 292 EPRFTYAGSSVKKYRPKYCGSCVDRCCTPLQTRTKKRFCEDEGEMFSKNVMIOSG 351

```

```

Db      264 KPFEELSGCTSMYTAKEGCTDRCCTPHRTTLLPVEFKCPDEP/MAKSMF1KTC 323
QY      352 KCNTNCPHNEASFRLY--SLFNDI 374
Db      324 ACHYNCPGDNDIFESLYRKMYGDM 348

RESULT 9
CTGF_PIG STANDARD; PRT; 349 AA.
ID CTGF_PIG
AC 019113;
DF 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Connective tissue growth factor precursor.
OS CTGF.
SN Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9623;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=uterus;
RX MEDLINE=97390475; PubMed=9242708;
RA Briggs R.A.; Steffen C.L.; Kim G.Y.; Vegunta R.K.; Diehl J.R.;
RA Harding P.A.;
RT Purification and characterization of novel heparin-binding growth
RT factors in uterine secretory fluids. Identification as heparin-
RT regulated Mr 10,000 forms of connective tissue growth factor."
RL J. Biol. Chem. 272:20275-20282(1997).
CC -1- FUNCTION: MAJOR CONNECTIVE TISSUE MITOATTRACTANT SECRETED BY
CC VASCULAR ENDOTHELIAL CELLS. PROMOTES PROLIFERATION AND
CC DIFFERENTIATION OF CHONDROCYTES (BY SIMILARITY). MEDIATES CELL
CC ADHESION AND ENHANCES FIBROBLAST GROWTH FACTOR-INDUCED DNA
CC SYNTHESIS (BY SIMILARITY).
CC -1- SUBUNIT: Monomer (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: FOUND IN THE EXTRACELLULAR MATRIX AND AS A
CC SOLUBLE FORM (BY SIMILARITY).
CC -1- SIMILARITY: Contains 1 IGFBP domain.
CC -1- SIMILARITY: Contains 1 WFPC domain.
CC -1- SIMILARITY: Contains 1 TSP type-1 domain.
CC -1- SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain.
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.1st-sib.ch/announce/
CC or send an email to license@1st-sib.ch).
CC -----
DR EMBL, U83916; AAC48756.1; -.
DR InterPro: IPR006208; Cys_knot.
DR InterPro: IPR006207; Cys_knot_C.
DR InterPro: IPR000867; Ins1_gro_fac_pr.
DR InterPro: IPR000884; TSP1.
DR InterPro: IPR001007; WFPC.
DR Pfam: PF00007; Cys_knot; 1.
DR Pfam: PF00219; IGFBP; 1.
DR Pfam: PF00090; TSP_1; 1.
DR Pfam: PF00093; WFPC; 1.
DR SMART: SM00041; CT; 1.
DR SMART: SM00121; IB; 1.
DR SMART: SM00209; TSP1; 1.
DR SMART: SM00214; WVC; 1.
DR PROSITE: PS01185; CTCK_1; 1.
DR PROSITE: PS01225; CTCK_2; 1.
DR PROSITE: PS00222; IGF_BINDING; 1.
DR PROSITE: PS00093; TSP1; 1.
DR PROSITE: PS01208; WFPC_1; 1.
DR PROSITE: PS01184; WVC_2; 1.
KW Cell adhesion; DNA synthesis; Extracellular matrix; Signal.

```

```

FT SIGNAL 1 26 POTENTIAL.
FT CHAIN 27 349 CONNECTIVE TISSUE GROWTH FACTOR.
FT DOMAIN 33 100 IGFBP.
FT DOMAIN 101 167 WFPC.
FT DOMAIN 198 243 TSP TYPE-1.
FT DOMAIN 256 330 CTCK.
FT DISULFID 256 293 BY SIMILARITY.
FT DISULFID 273 307 BY SIMILARITY.
FT DISULFID 284 323 BY SIMILARITY.
FT DISULFID 287 325 BY SIMILARITY.
FT DISULFID 292 329 BY SIMILARITY.
SO SEQUENCE 349 AA; 38007 MW; BB510E2B2B52D4A0 CRC64;

Query Match 43.7% Score 919; DB 1; Length 349;
Best Local Similarity 44.9%; Pred. No. 7.9e-61;
Matches 173; Conservative 62; Mismatches 102; Indels 48; Gaps 9;

QY 1 MSSSTFRTLLAVAVLTLHLTRIALSTCPAA-----CHCPL-EAPKCAPGVGLVRDGC 52
Db 1 MSATGLSPVACAFYLL-----LALCSRPASGGDCGCGCAAGKRACRACGATVLDGCG 56
QY 53 CKVCAKOLNEDCSKTQPCDHTKGLQCNFGASSTALKEICRAQSGRPECETNSRIYQNGES 112
Db 57 CRLCAKOLGELCTERDCDHPKGLFCDFGSPANKKICVCTAK-DGAPCVGEGTYRSGES 115
QY 113 FQPNCKHOCICIDAVGCIPLCPQELSLPNCGNPNLVKVSOGCCGEMWCDDESINDSL 172
Db 116 FQSSCKYQCTCLDAGVGVPCLSMDVRLPSPDCPPPRVVKLPGCCSEWVDEP--KDH- 172
QY 173 DDQDGLGLDASEVELRNNELIAGKSSLKRL-PVGEPEPVRLPNLHAHGOKCIYQT 231
Db 173 -----TYVGPALAVRLIEDTFGDPPTMM-----RANCLVQT 203
QY 232 TSMSCSKSGCTGISTVYNDNPECLVKEIRICEVRPGQPVYSSLKKGSKYTKSP 291
Db 204 TEMSACSKTCGCMGISTVYNDNAPCRLEKOSRLCMVBPCEADLEENIKKGCIRTPKIS 263
QY 292 EPRVFYTAGSSVYKYPKYGCGSCVDRCCTPLOTFTVKKMFRCDEDEMSKNYMMIOSC 351
Db 264 KPVEELSGCTSMYTAKEGCTDRCCTPHRTTLLPVEFKCPDEP/MAKSMF1KTC 323
QY 352 KCNTNCPHNEASFRLY--SLFNDI 374
Db 324 ACHYNCPGDNDIFESLYRKMYGDM 348

RESULT 10
NOV_CHICK STANDARD; PRT; 351 AA.
ID NOV_CHICK
AC P28686;
DF 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE NOV protein precursor (Nephroblastoma overexpressed gene protein).
GN NOV.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Brown Leghorn;
RX MEDLINE=92107157; PubMed=1309586;
RA Joliet V.; Martiniere C.; Dambirine G.; Plasiart G.; Brisac M.;
RA Crochet J.; Perbal B.;
RT "Proviral rearrangements and overexpression of a new cellular gene
RT (nov) in myeloblastosis-associated virus type 1-induced
RT nephroblastomas".
RL Mol. Cell. Biol. 12:10-21(1992).
CC -1- FUNCTION: IMMEDIATE-EARLY PROTEIN LIKELY TO PLAY A ROLE IN CELL
CC GROWTH REGULATION. ITS OVEREXPRESSION IS ASSOCIATED WITH
CC TUMORIGENESIS AND EXPRESSION OF A N-TERMINAL-TRUNCATED VERSION

```

```

CC OF NOV GENE IN CHICKEN EMBRYONIC FIBROBLASTS (CEF) IS SUFFICIENT
CC TO INDUCE THE TRANSFORMATION OF CEF IN VITRO.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: BRAIN AND HEART, AND AT A LOWER LEVEL IN
CC MUSCLE AND INTESTINE, IN THE EMBRIO. LONG AND LESS SO IN BRAIN AND
CC SPLEEN, IN ADULT CHICKEN.
CC -1- DEVELOPMENTAL STAGE: MAV1-INDUCED NEPHROBLASTOMAS EXPRESS A HIGH
CC LEVEL OF NOV GENE WHOSE TRANSCRIPTION IS NORMALLY ARRESTED IN
CC ADULT KIDNEY.
CC -1- SIMILARITY: Contains 1 IGFBP domain.
CC -1- SIMILARITY: Contains 1 WFRC domain.
CC -1- SIMILARITY: Contains 1 TSP type-1 domain.
CC -1- SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X59284; CAA1975.1; -
CC PIR: S20078; S20078.
CC InterPro: IPR006208; Cys_knot.
CC InterPro: IPR006207; Cys_knot_C.
CC InterPro: IPR000867; Insl_gro_fac_pr.
CC InterPro: IPR000884; TSP1.
CC InterPro: IPR001007; WMF_C.
CC Pfam: PF00007; Cys_knot; 1.
CC Pfam: PF00219; IGFBP; 1.
CC Pfam: PF00090; TSP_1; 1.
CC Pfam: PF00093; vwc; 1.
CC SMART: SM00041; CT; 1.
CC SMART: SM00121; IB; 1.
CC SMART: SM00209; TSP1; 1.
CC SMART: SM00214; vwc; 1.
CC PROSITE: PS01185; CTCK_1; 1.
CC PROSITE: PS01225; CTCK_2; 1.
CC PROSITE: PS00222; IGF_BINDING; 1.
CC PROSITE: PS50092; TSP1; 1.
CC PROSITE: PS01208; WMF_1; 1.
CC PROSITE: PS50184; WMF_2; 1.
CC Proto-oncogene; Growth factor; Signal.
CC KW SIGNAL; 1 24 POTENTIAL.
CC FT CHAIN; 25 351 NOV PROTEIN.
CC FT DOMAIN; 31 103 IGFBP.
CC FT DOMAIN; 104 170 WMF.
CC FT DOMAIN; 201 246 TSP TYPE-1.
CC FT DOMAIN; 258 332 CTCK.
CC FT DISULFID; 258 295 BY SIMILARITY.
CC FT DISULFID; 275 309 BY SIMILARITY.
CC FT DISULFID; 286 325 BY SIMILARITY.
CC FT DISULFID; 289 327 BY SIMILARITY.
CC FT DISULFID; 294 331 BY SIMILARITY.
CC CARBOHYD; 274 274 N-LINKED (GLCNAC... ) (POTENTIAL).
CC SQ SEQUENCE 351 AA; 38268 MW; 1ECB3FA3058C6797 CRC64;
Query Match 40.98; Score 861; DB 1; Length 351;
Best Local Similarity 44.58; Pred. No. 1.5e-56;
Matches 165; Conservative 44; Mismatches 110; Indels 52; Gaps 7;

```

```

OY 176 DDLGLDASEVELTRNNELIAIGKSLRLPVFTEPRVLPPLAHGOKCIQVOTTS 235
D 188 EATLGLDVS-----SSANCIQOTTS 210
OY 236 QCSKSGTGISRTVNDNEPCRLVKEFRICEVPCGOPYSSLKGRKSKTSKSPYR 295
D 211 ACSKSGMGFSFTVTRNRNOQCEVKQTRLCMRPCENEPSPD-KKGKCIQTRKSMKAVR 269
OY 296 FTYAGSSYKRRPKKCGSCVNGRCCTPLOTFRVKRRFCGCEGEMSKVMVMSQCKAV 355
D 270 FEYKNTSVQTKPRYGLCNDGRCCTPHNTKIQVERPCDQSKFLKRPMLINTVCVG 329
OY 356 NCPHNEASFR 366
D 330 NCPQSNNAFFQ 340
RESULT 11
NOV_COTJA STANDARD; PRT; 353 AA.
AC P42642;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
GN NOV protein precursor (Nephroblastoma overexpressed gene protein).
OS Coturnix coturnix japonica (Japanese quail).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Coturnix.
OX NCBI_Taxid=93934;
RN [1]
RP SEQUENCE FROM N.A.
RA Weiskirchen R., Bister K.;
RL Submitted (Aug-1994) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: IMMEDIATE-EARLY PROTEIN LIKELY TO PLAY A ROLE IN CELL
CC GROWTH REGULATION (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Secreted (By similarity).
CC -1- SIMILARITY: Contains 1 IGFBP domain.
CC -1- SIMILARITY: Contains 1 WFRC domain.
CC -1- SIMILARITY: Contains 1 TSP type-1 domain.
CC -1- SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U13063; AAA2128.1; -
CC InterPro: IPR006208; Cys_knot.
CC InterPro: IPR006207; Cys_knot_C.
CC InterPro: IPR000867; Insl_gro_fac_pr.
CC InterPro: IPR000884; TSP1.
CC InterPro: IPR001007; WMF_C.
CC Pfam: PF00007; Cys_knot; 1.
CC Pfam: PF00219; IGFBP; 1.
CC Pfam: PF00090; TSP_1; 1.
CC Pfam: PF00093; vwc; 1.
CC SMART: SM00041; CT; 1.
CC SMART: SM00121; IB; 1.
CC SMART: SM00209; TSP1; 1.
CC SMART: SM00214; vwc; 1.
CC PROSITE: PS01185; CTCK_1; 1.
CC PROSITE: PS01225; CTCK_2; 1.
CC PROSITE: PS00222; IGF_BINDING; 1.
CC PROSITE: PS50092; TSP1; 1.
CC PROSITE: PS01208; WMF_1; 1.
CC PROSITE: PS50184; WMF_2; 1.
CC Proto-oncogene; Growth factor; Signal.
KW

```

FT	SIGNAL		1.	26	POTENTIAL.
FT	CHAIN		27	353	NOV PROTEIN.
FT	DOMAIN		33	105	IGFBP.
FT	DOMAIN		106	172	WFCC.
FT	DOMAIN		203	248	TSP TYPE-1.
FT	DOMAIN		260	334	CTCK.
FT	DISULFID		260	297	BY SIMILARITY.
FT	DISULFID		277	311	BY SIMILARITY.
FT	DISULFID		288	327	BY SIMILARITY.
FT	DISULFID		291	329	BY SIMILARITY.
FT	DISULFID		296	333	BY SIMILARITY.
FT	CARBOHYD		276		N-LINKED (GLCNAC).
SO	SEQUENCE		353 AA:	38667 MM;	71DP9F8533882E89 CRC64; .

Query Match	40.7%;	Score 856;	DB 1;	Length 353;
Best Local Similarity	43.6%;	Pred. NO. 3.5e-56;		
Matches 163;	Conservative 47;	Mismatches 108;	Indels 56;	Gaps 7

QY	10	A V A V T L H L F R I A L-----S Y C P A A C--H C P L E A P K C A P G V G L V A D G C G C C Y C A K	58
Db	8	S L P U L L L L L L L L L R P S E V N G R E A C P R P C G R C A E P R C A P V P A V A D G C G C C L Y C A R	67
QY	59	Q L A N E D C S K T Q P O D H T K G L E C N F G A S T A L K G I C R A O S B G R P C E Y N S R I P O N G S E P O P N C K	118
Db	68	O R G S C S P L P C D S G S G L Y C D R G P E D G G T G I C M V -L E G D N C V F D G M I T R N E T Q P O S K	126
QY	119	H O C H T C I D A V A N C I P L C P O E L S L P N I G C N P R L Y K V S G O C C E W C D -E D S I-----K D S L	172
Db	127	Y O C T C R D P O J O L C P R C N I G L L L P G D C P F P K R I E V P G C E K W C E P R D E V L L G F A M A A	186
QY	173	D D O O D L I G L D A S E V E L T R N N E L I A I G K S S L K R L P V F T E P R V L F N P L H A H Q K C I V O T	232
Db	187	Y R O E P T A I D I V S D-----S A N C I E O T	209
QY	233	S M S O C S K S C T G I S T R V T N D N P E C R L Y K E T R I C E V R P C G O P Y S S L K R 3 K K C S K T K S P E	292
Db	210	E M S A C S R C G G E S P R V T N R N O Q C E M V A Q O T R C K M M R P C E N E R S P D -K K G K K C I P T R K S M K	268
QY	293	P V R P T Y A C C S S V K K I P R K Y I G C S C V D R C C T P L O T F Y V M R P R C E D G E M F S K N V M I O S K	352
Db	269	A V R E Y K N C T S V Q Y T P R Y C G L C N D R C C T P H N T K T I Q V E F R C P O G K F L K R P M L I N T C V	328
QY	353	C N Y N C P H P N E A S F R	366
Db	329	C H G N C P O S N A F F Q	342

RESULT 12	
NOV_HUMAN	STANDARD: PRT: 357 AA.
ID	NOV_HUMAN
AC	P46745; Q96BY5;
DT	01-FEB-1996 (Rel. 33, Created)
DT	01-FEB-1996 (Rel. 33, Last sequence update)
DE	15-SEP-2003 (Rel. 42, Last annotation update)
DN	NOV protein homolog precursor (NOV) (Nephroblastoma overexpressed gene protein homolog).
GN	NOV.
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Homidae; Hom.
OX	KCBI_TaxID=9606;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	TISSUE=Placenta;
RX	MEDLINE=94336229; PubMed=7520150;
RA	Mattnerie C., Huff V., Joubert I., Badeloch M., Saundere G.F.,
RT	"Structural analysis of the human nov proto-oncogene and expression in Wilms tumor";
RL	Oncogene 9:2729-2732(1994).
RN	[2]
RP	SEQUENCE FROM N.A.
RA	Jiang D., Gou D., Li W.;

DR SMART; SM00121; IB; 1.
 DR SMART; SM00209; TSP1; 1.
 DR SMART; SM00214; VWC; 1.
 DR PROSITE; PS01185; CTCK_1; 1.
 DR PROSITE; PS01225; CTCK_2; 1.
 DR PROSITE; PS00222; IGF_BINDING; 1.
 DR PROSITE; PS50092; TSP1; 1.
 DR PROSITE; PS01208; VWC_1; 1.
 DR PROSITE; PS50184; VWC_2; 1.
 KM Growth factor; Signal.
 FT SIGNAL 1 27
 FT CHAIN 28 357
 FT DOMAIN 35 107
 FT DOMAIN 108 174
 FT DOMAIN 205 250
 FT DISULFID 264 338
 FT DISULFID 264 301
 FT DISULFID 281 315
 FT DISULFID 292 331
 FT DISULFID 295 333
 FT DISULFID 300 337
 FT CARBOHYD 97 97
 FT CARBOHYD 280 280
 FT CONFLICT 97 97
 SQ SEQUENCE 357 AA; 39162 MW; 035D5BF4576B85B CRC64;

Query Match 39.9%; Score 839.5; DB 1; Length 357;
 Best Local Similarity 43.8%; Pred. No. 5.8e-55;
 Matches 163; Conservative 52; Mismatches 116; Indels 41; Gaps 9;

QY 1 MSSTF-----RTLAVALTLHLTRALST--CPAAC--HCPLEAPKAPGVGLVRCGCG 52
 DB 4 VOSTSECLRKQCLCLFELHLHLLGQVAATQRCPPGRCPPAPPPCAPGVRAVLDCGSC 63
 QY 53 CKVCAQOLNEDCSKTQPCDHTKGLFCNFGASSPALGICRAQSEGRPCENSRITYNQGES 112
 DB 64 CLVCAQORGECSGLDEPCDESSGLYCDRSDPNSGTGICLA-VEGDNCFYDGYITSGEK 122
 QY 113 FQPNCKHQCTCIDGAVGCIPLCPQELSLPFLGCPNRLKVSQCCCEWVCDSDITKSL 172
 DB 123 FQPSCKFQCTCRDQIGCVPRCQDVLLEPPNCPARKVEVPECECKMFCDE-EDSL 181
 QY 173 DDQDDLGLDASVEVLTNNELIATKGSGLKRLPVFGEPRLVNLPLAHGCKIVQTT 232
 DB 182 GG-----LTLAAYPEATLGEV-----SDSSV-----NCTEQTT 211
 QY 233 SMSGCKSGCTGISTRTVNDNPECRVYKERICEVAPCGQ-PYSSLLKKGKCKSKTKSP 291
 DB 212 EMTACKSGCMGFSTVYTNRRNKCEMLKOTRLQCMVRCDEPEQPTDKKCKLRTKSL 271
 QY 292 EPRVFTYAGCSYKRYPRYCGSCVDGRCCTPLQTRTVKMRFCEDGEMFSKNVMMIQSC 351
 DB 272 KAIHLQFKNCTSLHYTKPRFCGVCSDGRCCTPHNTTIAEPQCSGQIVKRVMMIGTC 331
 QY 352 KCVNNGPHRNEA 363
 DB 332 TCFNCPKNEA 343

RESULT 13
 NOV_RAT
 ID NOV_RAT STANDARD; PRT; 351 AA.
 AC 09Q0Z05;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE NOV protein homolog precursor (Novb) (Nephroblastoma overexpressed gene protein homolog).
 DE gene protein homolog).
 GN NOV.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 NCBI_TaxID=10116;

RN [1]
 RP SEQUENCE FROM N.A. AND CHARACTERIZATION.
 RC STRAIN-Sprague-Dawley;
 RX MEDLINE=20035752; Pubmed=10570975;
 RA Liu C., Liu X.J., Crowe P.D., Keiner G.S., Fan J., Barry G., Mann F.,
 RA Ling N., De Souza E.B., Maki R.A.;
 RT "Nephroblastoma overexpressed gene (NOV) codes for a growth factor
 RT that induces protein tyrosine phosphorylation."
 RL Gene 238:471-478(1999).
 CC - FUNCTION: Can act as a growth factor for some cells and binds to a
 CC specific receptor that leads to the phosphorylation of a 221 kDa
 CC protein.
 CC - SUBUNIT: Interacts with FBLN1 (By similarity).
 CC - SUBCELLULAR LOCATION: Secreted.
 CC - TISSUE SPECIFICITY: Ubiquitous.
 CC - SIMILARITY: Contains 1 IGF1R domain.
 CC - SIMILARITY: Contains 1 TSP type-1 domain.
 CC - SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL: AF171936; AAD49371.1; -
 DR InterPro: IPR006208; Cys_knot.
 DR InterPro: IPR006207; Cys_knot_C.
 DR InterPro: IPR000867; Ins1_gro_fac_pr.
 DR InterPro: IPR000884; TSP1.
 DR InterPro: IPR001007; VWC_C.
 DR Pfam: PF000007; Cys_knot; 1.
 DR Pfam: PF000219; IGFBP; 1.
 DR Pfam: PF000090; TSP_1; 1.
 DR Pfam: PF000093; VWC; 1.
 DR SMART; SM00041; CT; 1.
 DR SMART; SM00121; IB; 1.
 DR SMART; SM00209; TSP1; 1.
 DR SMART; SM00214; VWC; 1.
 DR PROSITE; PS01185; CTCK_1; 1.
 DR PROSITE; PS01225; CTCK_2; 1.
 DR PROSITE; PS00222; IGF_BINDING; 1.
 DR PROSITE; PS50092; TSP1; 1.
 DR PROSITE; PS01208; VWC_1; 1.
 DR PROSITE; PS50184; VWC_2; 1.
 KW Growth factor; Signal.
 FT SIGNAL 1 21
 FT CHAIN 22 351
 FT DOMAIN 29 101
 FT DOMAIN 102 168
 FT DOMAIN 199 244
 FT DOMAIN 268 332
 FT DISULFID 258 295
 FT DISULFID 275 309
 FT DISULFID 286 325
 FT DISULFID 289 327
 FT DISULFID 294 331
 FT CARBOHYD 91 91
 FT CARBOHYD 274 274
 SQ SEQUENCE 351 AA; 38509 MW; 02619707DE7C1BFB CRC64;

Query Match 39.4%; Score 829.5; DB 1; Length 351;
 Best Local Similarity 43.9%; Pred. No. 3.1e-54;
 Matches 158; Conservative 51; Mismatches 114; Indels 37; Gaps 8;
 QY 9 LAVAVTLHLTRALST--CPAAC--HCPLEAPKAPGVGLVRCGCGCKVCAQOLNEDC 64
 DB 10 LCLGFLHLHLNLQVATLNLCPRCPSQCSISPTCAPGVRAVLDCSCCPVCAQORGECS 69
 QY 65 SKTPQCDHTKGLFCNFGASSPALGICRAQSEGRPCENSRITYNQGESFQPNCKHQCTCI 124

```

Db 70 SEMRPPDQSSGLVCDRSADPNNGEIGCMV-PEGDNCVFDGVIYRNGEKI-EPNQCQYHCTCR 128
QY 125 DGAVGCTPLCPQELSLPNLGCNPNRLKVSAGCCCEEWVCEDSKSLDDQDLGLGDAS 184
Db 129 DGOIGCVPRQQLVLLPDPCCPAKKAAYVGECEKWTGCEE-KGTLAIG---LALPAY 183
QY 185 EVELTRNNELIAGKSSLRKLPVFGTEPRVLEFNPLAHGOKCIYVOTISQSCSKSGCTG 244
Db 184 REPRATVGEV-----SDSSI-----NCIEQTTIMKASCSKSGCMG 217
QY 245 ISTRVYNDNPECRVLETRICEVRPCGQ-PVYSLLKKGKSKTKSPFEVRYAGCSS 303
Db 218 LSTRVYNDNPECRVLETRICEVRPCGQ-PVYSLLKKGKSKTKSPFEVRYAGCSS 277
QY 304 VKYRRKRYGSCVDGRCCTPLQTRVYKMRRCEDGEFESKNVMIMIOSCICNYMCPHPNEA 363
Db 278 LTYRKPRFCIGICSDGRCPFNKTIQVEFQCLPQGIKKRPVAVIGTCVCHSNCPONNEA 337

RESULT 14
NOV_MOUSE STANDARD; PRT; 354 AA.
ID NOV_MOUSE
AC 064239;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE NOV protein homolog precursor (NOVH) (Nephroblastoma overexpressed
  gene protein homolog).
GN NOV.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
  SEQUENCE FROM N.A.
  STRAIN=129/SV, and ICR; TISSUE=Brain;
  MEDLINE=97131523; PubMed=8975721;
  RA Smith M.R., Natarajan D., Taylor L.B., Choi C.P., Martinale C.,
  RA Perbal B., Schofield P.N., Boulter C.A.;
  RT "Genomic structure and chromosomal mapping of the mouse nov gene.";
  RL Genomics 38:425-428(1996).
  [2]
  SEQUENCE FROM N.A.
  STRAIN=C57BL/6;
  MEDLINE=96204003; PubMed=8622864;
  RA Martinale C., Chevallier G., Rauscher F.J. III, Perbal B.;
  RT "Regulation of nov by WT1: a potential role for nov in
  nephrogenesis.";
  RL Oncogene 12:1479-1492(1996).
  CC -|- FUNCTION: IMMEDIATE-EARLY PROTEIN LIKELY TO PLAY A ROLE IN CELL
  CC GROWTH REGULATION (BY SIMILARITY).
  CC -|- SUBUNIT: Interacts with FBLN1 (By similarity).
  CC -|- SUBCELLULAR LOCATION: Secreted.
  CC -|- SIMILARITY: Contains 1 IGFBP domain.
  CC -|- SIMILARITY: Contains 1 WFCC domain.
  CC -|- SIMILARITY: Contains 1 TSP type-1 domain.
  CC -|- SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain.
  CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
  CC the European Bioinformatics Institute. There are no restrictions on its
  CC use by non-profit institutions as long as its content is in no way
  CC modified and this statement is not removed. Usage by and for commercial
  CC entities requires a license agreement (See http://www.1st-sib.ch/announce/
  CC or send an email to license@1st-sib.ch).
  CC EMBL: X97863; CAA66457.1; -
  DR EMBL: Y09257; CAA70454.1; -
  DR EMBL: X96585; CAA65404.1; -
  DR MGD: MGI:109185; NOV.
  DR InterPro: IPR006208; Cys_knot.
  DR InterPro: IPR006207; Cys_knot_C.

```

```

DR InterPro: IPR000867; Ins1_gro_fac-pr.
DR InterPro: IPR000884; TSP1.
DR InterPro: IPR001007; WFCC.
DR Pfam: PF00007; Cys_knot; 1.
DR Pfam: PF00219; IGFBP; 1.
DR Pfam: PF00090; TSP_1; 1.
DR Pfam: PF00093; WFC; 1.
DR SMART: SM00041; CT; 1.
DR SMART: SM00121; IB; 1.
DR SMART: SM00209; TSP1; 1.
DR SMART: SM00214; WFC; 1.
DR PROSITE: PS01185; CTCK_1; 1.
DR PROSITE: PS01225; CTCK_2; 1.
DR PROSITE: PS00223; IGF_BINDING; 1.
DR PROSITE: PS00092; TSP1; 1.
DR PROSITE: PS01208; WFC_1; 1.
DR PROSITE: PS50184; WFC_2; 1.
DR Growth factor; signal.
KW SIGNAL
FT SIGNAL 1 21
FT CHAIN 22 354
FT DOMAIN 29 101
FT DOMAIN 102 168
FT DOMAIN 202 247
FT DOMAIN 261 335
FT DISULFID 261 298
FT DISULFID 278 312
FT DISULFID 289 328
FT DISULFID 292 330
FT DISULFID 297 334
FT CARBOHYD 91 91
FT CARBOHYD 277 277
SQ SEQUENCE 354 AA; 38928 MW; 08EC8C6C67829DE CRC64;

Query Match 39.38; Score 827; DB 1; Length 354;
Best Local Similarity 44.28; Pred. NO. 4.8e-54;
Matches 161; Conservative 51; Mismatches 114; Indels 38; Gaps 9;

QY 7 RTLAVALNTLH-ITRLALS-TCPRAC--HGPLEAPKAPGVGLRDGCGCKYAKOLNE 62
Db 8 KCLGLGLPLHLISQVASLRCPSPKCPSPISPTCAPGVRSYLDGSCCPVCAKORGE 67
QY 63 DCSKTOPCDHTKGLIECNFGASSTALNGICRAOSEGRPCENSRITYQNGESFQPNCKHQT 122
Db 68 SCSEMRPPDQSSGLVCDRSADPNNGEIGCMV-PEGDNCVFDGVIYRNGEKI-EPNQCQYHCTCR 126
QY 123 CIDGAVGCTPLCPQELSLPNLGCNPNRLKVSAGCCCEEWVC--DEDSIKSLDDQDLGLG 180
Db 127 CRDQIGICLPRCQLDVLPPDPCCPAKKAAYVGECEKWTGSDGCTGTLG---LA 182
QY 181 LDASEVLETRNNELIAGKSSLRKLPVFGTEPRVLEFNPLAHGOKCIYVOTISQSCSKS 240
Db 183 LPAYRPRATVGEV-----SDSSI-----NCIEQTTIMKASCSKSGCMG 216
QY 241 CGGISSTRVYNDNPECRVLETRICEVRPCGQ-PVYSLLKKGKSKTKSPFEVRYAGCSS 299
Db 217 CGAGVSTRVYNDNPECRVLETRICEVRPCGQ-PVYSLLKKGKSKTKSPFEVRYAGCSS 276
QY 300 GCSVYKRYKRYGSCVDGRCCTPLQTRVYKMRRCEDGEFESKNVMIMIOSCICNYMCPHPNEA 359
Db 277 NCTSLYKPRFCIGICSDGRCPFNKTIQVEFQCLPQGIKKRPVAVIGTCVCHSNCPONNEA 336
QY 360 PNEA 363
Db 337 NNEA 340

RESULT 15
NOV_XENLA STANDARD; PRT; 343 AA.
ID NOV_XENLA
AC P51609;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 5, 2003, 14:02:31 ; Search time 32.913? Seconds

(without alignments)
2971.515 Million cell updates/sec

Title: US-09-495-448a-2

Perfect score: 2103
Sequence: 1 MSSSTFRTLAVALTLHLTR.....PNEASPRLYSLFNDIHKFRD 379

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Listing first 45 summaries

```

1: SPREMBL_23:*
2: sp.archaea:*
3: sp.bacteria:*
4: sp.fungi:*
5: sp.human:*
6: sp.invertebrate:*
7: sp.mammal:*
8: sp.mhc:*
9: sp.organelle:*
10: sp.phage:*
11: sp.plant:*
12: sp.podent:*
13: sp.virus:*
14: sp.vertebrate:*
15: sp.virus:*
16: sp.bacteriophage:*
17: sp.archaea:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2076	98.7	379	11 Q9WTM9	Q5TM9 ratius norv
2	1624.5	77.2	375	13 Q98TX5	Q58TX5 xenopus lae
3	1455	69.2	334	4 Q9UID7	Q9UID7 homo sapien
4	938	44.6	344	13 Q98T08	Q58T08 gallus gall
5	937	44.6	343	13 Q42607	Q42607 xenopus lae
6	920.5	43.8	347	13 Q9PT80	Q5PT80 notophthalm
7	891	42.4	349	6 Q97765	Q97765 sus scrofa
8	827	39.3	354	11 Q8CA67	Q8CA67 mus musculu
9	773	36.8	367	4 Q95388	Q95388 homo sapien
10	758.5	36.1	367	11 Q54775	Q54775 mus musculu
11	757	36.0	367	11 Q99P00	Q99P00 ratius norv
12	635.5	30.2	331	4 Q95958	Q95958 homo sapien
13	622	29.6	354	4 Q95389	Q95389 homo sapien
14	518.5	24.7	251	11 Q8CIC8	Q8CIC8 mus musculu
15	499.5	23.8	280	4 Q9HCS3	Q9HCS3 homo sapien
16	495.5	23.6	176	13 Q9PSS6	Q9PSS6 gallus gall

17	374.5	17.8	128	11 Q9R2C0	Q9R2C0 ratius norv
18	345	16.4	113	11 Q92164	Q92164 ratius norv
19	343	16.3	119	11 Q91V29	Q91V29 mus musculu
20	343	16.3	119	11 Q920W6	Q920W6 mus spicile
21	341	16.2	100	4 Q9UDL6	Q9UDL6 homo sapien
22	298.5	14.2	374	5 Q9VVR0	Q9VVR0 drosophila
23	296	14.1	470	5 Q9VVK3	Q9VVK3 drosophila
24	293	13.9	230	4 Q8VYK7	Q8VYK7 homo sapien
25	195	9.3	77	4 Q9UDE4	Q9UDE4 homo sapien
26	184	8.7	70	13 Q9DF21	Q9DF21 scyllorhinu
27	179	8.5	2327	13 Q9IBG7	Q9IBG7 xenopus lae
28	171	8.1	1045	5 Q8T3A6	Q8T3A6 caenorhabdi
29	171	8.1	1070	5 Q8T3A7	Q8T3A7 caenorhabdi
30	171	8.1	1111	5 Q9XMD6	Q9XMD6 caenorhabdi
31	167.5	8.0	792	13 Q90Z43	Q90Z43 gallus gall
32	167.5	8.0	1095	13 Q90XG4	Q90XG4 gallus gall
33	167	7.9	58	6 Q97574	Q97574 bos taurus
34	166	7.9	1034	11 Q35888	Q35888 ratius norv
35	166	7.9	1521	4 Q95710	Q95710 homo sapien
36	166	7.9	1525	4 Q9Y5Q7	Q9Y5Q7 homo sapien
37	166	7.9	1529	4 Q94813	Q94813 homo sapien
38	165.5	7.9	950	13 Q90Z44	Q90Z44 gallus gall
39	163	7.8	1036	4 Q9NZV1	Q9NZV1 homo sapien
40	162	7.7	360	5 Q8MN64	Q8MN64 dictyosteli
41	161.5	7.7	406	11 Q92513	Q92513 mus musculu
42	161.5	7.7	426	11 Q8VBA6	Q8VBA6 mus musculu
43	160.5	7.6	2104	5 Q21281	Q21281 caenorhabdi
44	160.5	7.6	2104	5 Q964N4	Q964N4 caenorhabdi
45	158.5	7.5	1028	11 Q9JLL0	Q9JLL0 mus musculu

ALIGNMENTS

RESULT 1

ID	Q9WTM9	PRELIMINARY:	PRT:	379 AA.
AC	Q9WTM9	01-NOV-1999 (TRENBLREL. 12, Created)		
DT	01-NOV-1999 (TRENBLREL. 12, Last sequence update)			
DE	01-MAR-2003 (TRENBLREL. 23, Last annotation update)			
DE	CYR61 precursor.			
GN	CYR61.			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
OX	NCBI_TaxID=10116;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
KC	STRAIN=Izm; TISSUE=Aorta;			
RA	Unoki H., Yonekura H., Furukawa K., Yamamoto H.;			
RT	"Rat Cyr61 mRNA."			
RL	Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.			
DR	EMBL: AB015877; BAA78339.1; -			
DR	InterPro: IPR006208; Cys_knot.			
DR	InterPro: IPR006207; Cys_knot_C.			
DR	InterPro: IPR000867; Insl_gro_fac_pr.			
DR	InterPro: IPR000884; TSPL.			
DR	InterPro: IPR001007; VWF_C.			
DR	Pfam: PF00007; Cys_knot; 1.			
DR	Pfam: PF00219; IGFBP; 1.			
DR	Pfam: PF00090; tsp_1; 1.			
DR	Pfam: PF00093; vwc; 1.			
DR	SMART: SM00041; CT; 1.			
DR	SMART: SM00121; IB; 1.			
DR	SMART: SM00209; TSPL; 1.			
DR	SMART: SM00214; VWC; 1.			
DR	PROSITE: PS01185; CTCK_1; 1.			
DR	PROSITE: PS01225; CTCK_2; 1.			
DR	PROSITE: PS00222; IGF_BINDING; 1.			
DR	PROSITE: PS00092; TSPL; 1.			
DR	PROSITE: PS01208; VWF_C; 1.			
KW	Signal.			

FT SIGNAL 1 24 POTENTIAL.
 CHAIN 25 379 CYR61.
 SO SEQUENCE 379 AA: 41728 MW: D2ABAF077B84762B CRC64:

Query Match 98.7%; Score 2076; DB 11; Length 379;
 Best Local Similarity 98.2%; Pred. No. 1.8e-189;
 Matches 372; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 MSSSTRATAVANTLTLTRALSTCPACCHCPLEAPKCAPGVGLVDRDGGCCCKVCAKOL 60
 DB 1 MSSSTRATAVANTLTLTRALSTCPACCHCPLEAPKCAPGVGLVDRDGGCCCKVCAKOL 60
 QY 61 NEDCSSTQPCDHRTKGLCECNFASSTALKGICRAQSEGRPEYNSRIYQNGESFQPNCKHQ 120
 DB 61 NEDCSSTQPCDHRTKGLCECNFASSTALKGICRAQSEGRPEYNSRIYQNGESFQPNCKHQ 120
 QY 121 CTCIDGAVGICPLCPPELSPNLCGPNPLVYKSGCCCEWVCDDEDSIKDSLDDDDLLG 180
 DB 121 CTCIDGAVGICPLCPPELSPNLCGPNPLVYKSGCCCEWVCDDEDSIKDSLDDDDLLG 180
 QY 181 LDASEVELTRNNELIAGKSSSLKRLPVEGTEPRVLFNPLHAHGQKCIYQTTSMSCSKS 240
 DB 181 LDASEVELTRNNELIAGKSSSLKRLPVEGTEPRVLFNPLHAHGQKCIYQTTSMSCSKS 240
 QY 241 CGTGISTRTVNDNPECHLVKTRICVRRPGGQPVYSSLLKGGKSKTKKSPPEVRYTYAG 300
 DB 241 CGTGISTRTVNDNPECHLVKTRICVRRPGGQPVYSSLLKGGKSKTKKSPPEVRYTYAG 300
 QY 301 GSSVKYRPRKYGSCVDGRCCTPIQRTYKMRFCEDGEMFSKNVMIIOSCKNVCNCPHP 360
 DB 301 GSSVKYRPRKYGSCVDGRCCTPIQRTYKMRFCEDGEMFSKNVMIIOSCKNVCNCPHP 360
 QY 361 NEASFRLYSLFNDIHKFRD 379
 DB 361 NEASFRLYSLFNDIHKFRD 379

RESULT 2

Q98TX5 PRELIMINARY; PRT; 375 AA.
 AC 098TX5;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Secreted cysteine-rich protein cyr61.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipiloidea; Pipidae;
 OC Xenopodidae; Xenopus.
 NCBI_TaxID=6355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Latinkic B.V., Bennett B., Smith J.C.;
 RT "Characterization of Xenopus cyr61."
 RT Submitted (NOV-2000) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AF20592; AAK0947.1; -
 DR InterPro: IPR006208; Cys_knot.
 DR InterPro: IPR006207; Cys_knot.C.
 DR InterPro: IPR000867; Insl_gro_fac.
 DR InterPro: IPR000884; TSP1.
 DR InterPro: IPR001007; VWF.C.
 DR Pfam: PF00007; Cys_knot.1.
 DR Pfam: PF00090; IGFBP.1.
 DR Pfam: PF00093; vwc.1.
 DR SMART: SM00041; CT.1.
 DR SMART: SM00121; IB.1.
 DR SMART: SM00209; TSP1.1.
 DR SMART: SM00214; VWC.1.
 DR PROSITE: PS01185; CTCK.1; 1.
 DR PROSITE: PS01225; CTCK.2; 1.
 DR PROSITE: PS00222; IGF_BINDING.1.
 DR PROSITE: PS50092; TSP1.1.

DR PROSITE: PS01208; VWF.C.1.
 SO SEQUENCE 375 AA: 41460 MW: 78075CA7B380304E CRC64:

Query Match 77.2%; Score 1624.5; DB 13; Length 375;
 Best Local Similarity 77.5%; Pred. No. 1.9e-146;
 Matches 293; Conservative 29; Mismatches 37; Indels 19; Gaps 7;

QY 9 LAVALTLHLTRALSTCPACCHCPLEAPKCAPGVGLVDRDGGCCCKVCAKOLNEDCSKTQ 68
 DB 10 LAVAL-TSGFIDLAVSSCPAVGOCPEYKCAPGVGLVDRDGGCCCKVCAKOLNEDCSKTQ 68
 QY 69 PCDHRTKGLCECNFASSTALKGICRAQSEGRPEYNSRIYQNGESFQPNCKHCTCIDGAV 128
 DB 69 PCDHRTKGLCECNFASSTALKGICRAQSEGRPEYNSRIYQNGESFQPNCKHCTCIDGAV 128
 QY 129 GCPLCPPELSPNLCGPNPLVYKSGCCCEWVCDDEDSIKDSLDDDDLLG 184
 DB 129 GCPLCPPELSPNLCGPNPLVYKSGCCCEWVCDDEDSIKDSLDDDDLLG 184
 QY 185 EVELTRNNELIAGKSSSLKRLPVEGTEPRVLFNPLHAH--GQKCIYQTTSMSCSKS 241
 DB 187 EVELTRNNELIAGKSSSLKRLPVEGTEPRVLFNPLHAH--GQKCIYQTTSMSCSKS 241
 QY 242 GTGISTRTVNDNPECHLVKTRICVRRPGGQPVYSSLLKGGKSKTKKSPPEVRYTYAG 301
 DB 239 GTGISTRTVNDNPECHLVKTRICVRRPGGQPVYSSLLKGGKSKTKKSPPEVRYTYAG 301
 QY 302 GSSVKYRPRKYGSCVDGRCCTPIQRTYKMRFCEDGEMFSKNVMIIOSCKNVCNCPHP 361
 DB 299 GSSVKYRPRKYGSCVDGRCCTPIQRTYKMRFCEDGEMFSKNVMIIOSCKNVCNCPHP 361
 QY 362 EASFRLYSLFNDIHKFRD 379
 DB 359 EASFRLYSLFNDIHKFRD 375

RESULT 3

Q9UID7 PRELIMINARY; PRT; 334 AA.
 AC 09UID7;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Cyr61 protein.
 GN CYR61.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE OF 1-107 FROM N.A.
 RA Anding B., Long Y.;
 RT "Cloning of a new gene down-regulated in the small-cell tumor
 RT embryonal-rhabdomyosarcoma (RMS)."
 RT Submitted (MAY-1997) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AF003114; AAF21597.1; -
 DR InterPro: IPR006208; Cys_knot.
 DR InterPro: IPR006207; Cys_knot.C.
 DR InterPro: IPR000867; Insl_gro_fac.
 DR InterPro: IPR000884; TSP1.
 DR InterPro: IPR001007; VWF.C.
 DR Pfam: PF00007; Cys_knot.1.
 DR Pfam: PF00090; IGFBP.1.
 DR Pfam: PF00093; vwc.1.
 DR SMART: SM00041; CT.1.
 DR SMART: SM00209; TSP1.1.
 DR SMART: SM00214; VWC.1.
 DR PROSITE: PS01185; CTCK.1; 1.
 DR PROSITE: PS01225; CTCK.2; 1.
 DR PROSITE: PS01208; VWF.C.1.
 DR PROSITE: PS01208; VWF.C.1.
 SO SEQUENCE 334 AA: 37246 MW: 9188987A7352E948 CRC64:

Query Match	69.2%	Score 1455:	DB 4:	Length 3:4:
Best Local Similarity	78.9%	Pred. No. 2.3e-130:		
Matches 273:	Conservative 13:	MissMatches 32:	Indels 28:	Gaps 6
OY	49	GCGCCVCAKADLNEDCSKTOPCDHFKGEGECNFGASSTALKGICRAOSEIRPEYNSRIQ	108	
Db	2	GCGRHPNLCIHGHNTASPTKYKHHKKGLECNFGASTALKGICRAOSEIRPEYNSRIQ	61	
OY	109	NGESFQPNCKHOCTCIDGAVGCIPQLCPQELSLPNIGCPNRLVYKSGQICEBWCDSDSI	168	
Db	62	NGESFQPNCKHOCTCIDGAVGCIPQLCPQELSLPNIGCPNRLVYKSGQICEBWCDSDSI	121	
OY	169	KDSLIDODL-----IGLDAFEVLRNNNELIAGKSSLRPLPVGTE:RYVLPNPLHANG	224	
Db	122	KDPMDDQGLLAKETELGFDASFEVLRNNNELIAGKSSLRPLPVGME:RIYVNP.L--OG	179	
OY	225	QKCIYQTTWSQSCSKSCGCTGISTRTYNDNPECLRKETRIFICEVRPCQGIVYSLSLKGGKC	284	
Db	180	QKCIYQTTWSQSCSKCTGCTGISTRTYNDNPECLRKETRIFICEVRPCQGIVYSLSLKGGKC	239	
OY	285	SKTKSPSPVPRFTYAGSSVKKYRKYCGSCVDGKCCIPDQRTYVKMR:RCED-----	337	
Db	240	SKTKSPSPVPRFTYAGCLSVKKYRKYCGSCVDGKCCIPQLTRT-----CEDAVPLRRM	293	
OY	338	GGMSKN-----VMMIGCKMYNCPHPEASPRLSLTLNDHKKFD	379	
Db	294	GDITQERHDDPVL-----KQNTNCNCHAAEAPFYRLNDHKKFD	334	

ID	NAME	PROT	AA
098R08	PRELIMINARY;	344	AA.
AC	098R08.		
DT	01-JUN-2001 (TREMBLrel. 17, Created)		
DT	01-JUN-2001 (TREMBLrel. 17, Last sequence update)		
DE	01-MAR-2003 (TREMBLrel. 23, Last annotation update)		
DE	Connective tissue growth factor precursor (connective tissue growth factor/hypertrophic chondrocyte-specific protein 24).		
OC	Gallus gallus (Chicken).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Archosauaria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;		
OC	Gallus.		
OX	NCBI_taxid=9031;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RL	Gygi D.;		
RN	submitted (Jan-2001) to the EMBL/Genbank/DBJ databases.		
RP	[2]		
RA	SEQUENCE FROM N.A.		
RA	Mukudai Y., Kubota S., Takigawa M.;		
DR	Submitted (DEC-2001) to the EMBL/Genbank/DBJ databases.		
DR	EMBL; AJ298335; CAC33438.1; -		
DR	EMBL; AF463517; AL68834.1; -		
DR	InterPro: IPR006208; Cys_knot.		
DR	InterPro: IPR000867; Cys_knot.C.		
DR	InterPro: IPR000884; TSP1.		
DR	InterPro: IPR001007; WVF_C.		
DR	Pfam; PF00007; Cys_knot; 1.		
DR	Pfam; PF00219; IGFEBF; 1.		
DR	Pfam; PF00090; TSP_1; 1.		
DR	Pfam; PF00093; WVC; 1.		
DR	SMART; SM00041; CT; 1.		
DR	SMART; SM00121; IB; 1.		
DR	SMART; SM00209; TSP1; 1.		
DR	SMART; SM00214; WVC; 1.		
DR	PROSITE; PS01185; CMCK_1; 1.		
DR	PROSITE; PS01225; CMCK_2; 1.		
DR	PROSITE; PS00222; IGF_BINDING; 1.		
DR	PROSITE; PS00992; TSP1; 1.		
DR	PROSITE; PS01208; WVC; 1.		
FW	Signal.		

FT	SIGNAL	1	21	POTENTIAL.
FT	CHAIN	22	344	CONNECTIVE TISSUE GROWTH FACTOR.
SEQ	SEQUENCE	344 AA:	37499 MW:	69E639AF6BFD1D0 CRC64;
	Query Match	44.6%;	Score 938;	DB 13;
	Best Local Similarity	45.7%;	Pred. No. 44e-81;	
	Matches 170;	Conservative	61;	Mismatches 103; Indels 38; Gaps 7
QY		8 TLAAVAVTLHLH-TLALSTCPAACHCPL-EAKKCAPGVLYVDDGGCCCKVCAKQLNEDCS	65	
Db		5 SLAAVALLALALGPEVRGEGECGGCCGGSGPGSGCAGAVSLYVDDGGCCCKVCAKQNLGELCT	64	
QY		66 KTOPEDHTKGLTECNAGASTALGKICRAQSEGRPEEYNSRIYQNESFQPNKHQOCTID	125	
Db		65 ERDPCDHKKGLFCDGSGPANRRIRIGVCTAR-DEAPCVFSGSMYRSGESFSSGCKYQOTCTCD	123	
QY		126 GAVGCTPLCPQBELSLPNTGCPNRLPVKVSQGCCCEWVCDSDSIKSLDQDDLLGLDASE	185	
Db		124 GAVGCVPLCSMDVIRLPSDPCPYPRPVYKKLPKGCCEWVCDSE-----	163	
QY		186 VELTRNNNELIAGKSSSLKRL-PVGTGEPRVLFNPLHAHGKCIQVOTTSMSQCSKSGTG	244	
Db		164 -----AKQOTAVGPRLAAVRLIEDYTGPRPTM-----RANCLVOTITMSAKTCSGAG	211	
QY		245 ISTRVTDNPECRILVETRIICEVRPCGQPVYSSLAKGKKKSKTKRSPEDEVFTYAGCCSSV	304	
Db		212 ISTRVTNNAFCELRTEKOSRLCWVRCEADLEINIKKGGKCIIRTPKISPKIFELSGCTSV	271	
QY		305 KKYRKVYGGSCVDCGCTPIQTRVTKMRRCGDEGFNSNNMMIOSCKNNYNDPHNENAS	364	
Db		272 KTYRAKFGVGVCDGCKCTPHRTATLPVEKCDGELGEMKKMMKMFILTCACAHYNGGNDIT	331	

```

QY      305  KKRRKCYGSCDYGCTPLQRTYKMRPCEDEGEMFSKNVMKIOSCKNYNCPHNENAS 364
DB      272  KTYRAKFCGYCVDGRCCTPHRTATLPEVEKCPDGEIMKRKMEFKTCACAHYNCQNDJF 331
QY      365  FRLY--SLFNDI 374
DB      332  ESLYYRKMKGDM 343

RESULT 5
ID      042607      PRELIMINARY;      PRF;      343  AA:
AC      042607;
DT      01-JAN-1998  (TREMBLrel .05, Created)
DT      01-JAN-1998  (TREMBLrel .05, Last sequence update)
DT      01-MAR-1993  (TREMBLrel .23, Last annotation update)
DE      Connective tissue growth factor XTCGF.
OS      Xenopus laevis (African clawed frog).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;
OC      Xenopodinae; Xenopus.
CX      NCBI_TaxID=83355;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Ying X., King M.L.;
RL      Submitted (Aug-1997) to the EMBL/GenBank/DBJ databases.
DR      EMBL; U43524; AAB67639.1; .
DR      EMBL; U43523; AAB67638.1; .
DR      InterPro; IPR006208; Cys_knot.
DR      InterPro; IPR006207; Cys_knot_C.
DR      InterPro; IPR000867; Ins1_gro_fac_pr.
DR      InterPro; IPR000884; TSP1.
DR      InterPro; IPR001007; WVE_C.
DR      Pfam; PF00007; Cys_knot; 1.
DR      Pfam; PF00219; IGFBP; 1.
DR      Pfam; PF00090; tsp_1; 1.
DR      Pfam; PF00093; WVC; 1.
DR      SMART; SM00041; CT; 1.
DR      SMART; SM00121; IB; 1.
DR      SMART; SM00209; TSP1; 1.
DR      SMART; SMO0214; WVC; 1.
DR      PROSITE; PS01185; CMCK_1; 1.
DR      PROSITE; PS01225; CMCK_2; 1.
DR      PROSITE; PS50092; TSP1; 1.
DR      PROSITE; PS01208; WVEC; 1.

```

SEQ	SEQUENCE	343 AA;	37966 MW;	9382221C5DB565A81 CRC64;
Query Match	44.6%;	Score 937;	DB 13;	Length 343;
Best Local Similarity	46.7%;	Pred. No. 5.5e-81;		
Matches 165;	Conservative 60;	Mismatches 90;	Indels 38;	Gaps 6;
QY	CPACHCHCLEPDKCAPGVGLVRDGGCGGCKYCAKOLNEGCSKTPQCDHTKGLCECNAGASST	85		
DB	24 CNGEGCCGCKNKPVPDGVRYAMVDGCGCKKVCSSKQUGELCTEDVDYCDPHKGLFCDGSKYN	83		
QY	86 ALKGICRAOSEGRPCPEYNSIRYQNGESFQPNCKHOCTCIDGAVGCIPLCPQELSPNLGAC	145		
DB	84 KRIGCTAR-BGAPCVGCTGYRSGESFSQCKYCTCIDGVCVPLCSMDIRLPSEEC	142		
QY	146 PNPRLVYKSGGCGEWCWDEDSIKSLDDODDLG--LDASVELTRNNELAIKSSSL	203		
DB	143 PPRPVKLPGKCCSEWVCDQ-----PQERTLVGPAIPAFMEET-----	181		
QY	204 KRLPVGTEPRVLEPNLHAHGOKCIYQTTSMSCSKSGCTGISTREVTNDNEPCRLVKETR	263		
DB	182 -----YGDPSLI-----RANCLVQTTESMSACSKTGCMGISTRVTDNEHCRLKQSR	229		
QY	264 ICEVRPCGQPYSSLKRRKCKSKTKKSEPRFTATGSSVYKKYRPKTCGSCVDRCTCP	323		
DB	230 LCMVRPCPADLEENIKRKKKCIKPKISKPVKFEFSGGTIVYTKAKCGVCTDRCCTP	289		
QY	324 LQTFVKRFRFCEDEMFESKNVMMIOSCKVNCNCPHEDASPRLY--SLENDI	374		
DB	290 HRTALPVEFPCDGEVKKMMFTKTCACHFNGCGDNDIFEMVYRKMVYGDGM	342		
RESULT 6				
Q9PT80	PRELIMINARY;	PRT;	347 AA.	
ID	Q9PT80			
AC	Q9PT80;			
DT	01-MAY-2000 (TREMBLrel. 13, Created)			
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)			
DT	01-MAR-2003 (TREMBLrel. 23, Last annotation update)			
DE	Connective tissue growth factor.			
GN	CTGF.			
OS	Notophthalmus viridescens (Eastern newt) (Triturus viridescens).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Amphibia; Batrachia; Caudata; Salamandroides; Salamandridae;			
OC	Notophthalmus;			
OX	NCBI_TaxID=8316;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Forelimb blastema;			
RC	MEDLINE=99033008; PubMed=9813273;			
RA	Cash D.E., Gates P.B., Imokawa Y., Brookes J.P.;			
RA	"Identification of new connective tissue growth factor as a target of			
RL	retinoid regulation in limb blastenal cells.";			
RT	Gene 222,119-124(1998)			
RL	EMBL, AJ271167; CAB65965.1; -			
DR	InterPro: IPR006208; Cys_knot.			
DR	InterPro: IPR006207; Cys_knot.C.			
DR	InterPro: IPR000867; Insl_gro_fac_pr.			
DR	InterPro: IPR000884; TSP1.			
DR	InterPro: IPR001007; WVF_C.			
DR	Pfam: PF00007; Cys_knot; 1.			
DR	Pfam: PF00219; IGFBP; 1.			
DR	Pfam: PF00090; tsp_1; 1.			
DR	Pfam: PF00093; wvc; 1.			
DR	SMART: SM00041; CT; 1.			
DR	SMART: SM00121; IB; 1.			
DR	SMART: SM00209; TSP1; 1.			
DR	SMART: SM00214; WVC; 1.			
DR	PROSITE: PS01185; CTCK_1; 1.			

DR	PROSITE: PS01225; CKCK 2; 1.
DR	PROSITE: PS00222; IGF_BINDING; 1.
DR	PROSITE: PS50092; TSP1; 1.
DR	PROSITE: PS01208; VMFC; 1.
DR	SEQUENCE 347 AA; 38098 MW; 3B7E2399F27672C1 CRC64;
Query Match	43.8%; Score 920.5; DB 13; Length 347;
Best Local Similarity	45.1%; Pred. No. 2, 1e-79;
Matches 169;	Conservative 54; Mismatches 103; Indels 49; Gaps 6
QY	9 LAVALTLHLFRLALSTCPACHCPLLEAFKCAPGVGLVYDGGCGCCCKVCAKQLNEDCSKTQ 68
DB	12 LLLAVALLSWVSCA-QDCSGECRCRNKPEPCAGTSLVMDGCGCCCKVCAKQLGELCTEAD 70
QY	69 PODHRTKGLGNEGASTSLAKGTICRAOSEGRPCPEYNSIRYONSGEPQPNKHOCCTCIDGAV 128
DB	71 VCDPRHGLFCDGDSRYNKKIGVCTYK-DEAPCVFGGMYRSESFSSCKRYOCTCIDGAV 129
QY	129 GCIPPLCPDELSPNLGECPNRLVYKSGQCCSEWCEDEDSIKDSLDDQLGLDASEVEL 188
DB	130 GCVPLFGVDVRLSPDCEPPRRVVKQLGKCEEMVWCDDPK-----EQ 170
QY	169 TRNNELIA-----GKSSLKRLPVETEPRLVFNPLAHGQKCIYOTISMOSKSC 241
DB	171 TRVGPALAVYROEETVGPDSILMR-----ANCLVOTTEMSAKSKTC 211
QY	242 TGISTRTVYNDNPECRILVETRICERPCGQPVYSLKGGKCKSKKSPPEVRFETYAC 301
DB	212 GNGISTRTVYNDNEMCRLEKQSLCAWVRPEADLEENIKKCKICIRPKISKPVKPELSCG 271
QY	302 SSVKRYPRKYCGSVYDRCCTPLQTRTYVMKRFRCDEDEMFNSNNMAMIOSCKNCNYCPHPN 361
DB	272 TSVKRYPRKFCGVCVCDGRCCTPHRTATLPVEERKCPDGEYMKRKKMFIFKTCACHYNGPDN 331
QY	362 BASFRLY--SLFNDI 374
DB	332 DIFESMYRKMGIDM 346
RESULT 7	
ID	097765 PRELIMINARY; PRT; 349 AA.
AC	097765;
DT	01-MAY-1999 (TREMBLrel. 10, Created)
DT	01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT	01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE	Connective tissue growth factor.
GN	CTGF.
OS	Sus scrofa (Pig).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX	NCBI_TaxID=9823;
DR	NCBI
DR	RM
DR	RP
DR	SEQUENCE FROM N.A.
DR	RC
DR	TISSUE=uterus;
DR	Harding P.A.; Brigstock D.R.;
DR	"Cloning and sequencing of a porcine connective tissue growth factor
DR	(CTGF) cDNA."
DR	Submitted (SEP-1996) to the EMBL/Genbank/DBJ databases.
DR	EMBL; U70060; AAD00174.1; -
DR	InterPro: IPR006208; Cys_knot.
DR	InterPro: IPR006207; Cys_knot.C.
DR	InterPro: IPR000867; Insl_gro_fac_pr.
DR	InterPro: IPR000884; TSP1.
DR	InterPro: IPR001007; VMF_C.
DR	Pfam: PF00007; Cys_knot; 1.
DR	Pfam: PF00219; IGFBP; 1.
DR	Pfam: PF00090; tsp.1; 1.
DR	Pfam: PF00093; vwc; 1.
DR	SMART; SM00041; CT; 1.
DR	SMART; SM00121; IB; 1.
DR	SMART; SM00209; TSP1; 1.
DR	SMART; SM00214; VWC; 1.

DR PROSITE; PS00992; TSPL; 1.
DR PROSITE; PS01208; VWF; 1.
SO SEQUENCE 367 AA; 40331 MW; 9F29CA94D69C0502 CRC64;
Query Match 36.8%; Score 773; DB 4; Length 367;
Best Local Similarity 38.4%; Pred. No. 2,5e-65;
Matches 146; Conservative 58; Mismatches 124; Indels 52; Gaps 7;

9 LAVAVTLHLTRALST-----CPAACHCPLEAPKCAPGVGLVRDGGCCGCAKOLN 61
25 LSPAPITMDTPAPLEDTSSRFQCKWPCPCPPRCPGLVSLITDGCCECCMACQOLG 84
62 EDCSKTOPCDHTKLECNF--GASSTALKGICRAOSEGRPEYNSRIYONGESFQPNCKH 120
85 DMCTEALICDPHRLGYCDYSGDRPRYALIGVC-AQVVGVCVLGVRNNGSQSPNCKYN 143
121 CTCIDGAVGCIPLCPQELSLPNLNGCPRPLVYKSGGCCCEWVDEDSIKDSLDDDDL 180
144 CTCIDGAVGCTPLC-LRVRPRLMCPHPRVSIPIGHCCCEWVEDAKPRKTAPRDTGA 202
181 LDA-SEVELTRNNELIAGKSSLKRLPYGTEPRVLFNPLAHGCKCIQVOTSMSCSK 239
203 FDAVGEVEMHNR-----CIATYSPMSPCST 228
240 SCGTGISTRVNTDNPFCRLVETRICVAPCGQPVYSSLKGRKSKTKKSPVPRFTYA 299
229 SCGLGVSTRISNNACWPEQESRLCNLRPCVDVIDLITLAKGKCLAVYQPEASMFLLA 288
300 GCSVYKYPKYGCGSVDRCCPTLOTTRYKMRFBEDGEMFSKNMIMOSCKNCPH 359
289 GCISTSTYCPKYGVMDCNRCCLPYKSKTIDVSFCQDGLGFSROYLWLNACFCNLSCRN 348
360 PNEASFRLYSLFNDIHKFRD 379
349 PND-----IFADLESYPD 361

RESULT 10
054775 PRELIMINARY; PRT; 367 AA.
ID 054775;
AC 054775;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE ELM1.
GN ELM1 OR WISPL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=HEN;
RX MEDLINE=98119879; PubMed=9449709;
RA Hashimoto Y., Shindo-Okada N., Tani M., Nagamachi Y., Takeuchi K.,
RA Shiroishi T., Toma H., Yokota J.;
RT "Expression of the ELM1 gene, a novel gene of the CCN (connective
RT tissue growth factor, Cyr61/Ce10, and neuroblastoma overexpressed
RT gene) family, suppresses in vivo tumor growth and metastasis of K-1735
RT murine melanoma cells.";
RL J. Exp. Med. 187:289-296(1998).
RN (2)
RP SEQUENCE FROM N.A.
RC TISSUE=Mammary gland;
RX MEDLINE=99061933; PubMed=9843955;
RA Pennica D., Swanson T.A., Welsh J.W., Roy M.A., Lawrence D.A., Lee J.,
RA Brush J., Taneyhill L.A., Denel B., Lew M., Matanabe C., Cohen R.L.,
RA Melham M.F., Finley G.G., Quirke P., Goddard A.D., Hillan K.J.,
RA Gurney A.L., Botstein D., Levine A.J.;
RP Wisp genes are members of the connective tissue growth factor family
RT that are up-regulated in wt-1-transformed cells and aberrantly
RT expressed in human colon tumors.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:14717-14722(1998).

DR EMBL; AB004873; BAA24949.1; -
DR EMBL; AF100777; AAC96319.1; -
DR MGD; MGI:1197008; Wisp1.
DR InterPro; IPR006208; Cys_knot.
DR InterPro; IPR006207; Cys_knot_C.
DR InterPro; IPR000867; Insl_gro_fac_pr.
DR InterPro; IPR000884; TSPL.
DR InterPro; IPR001007; VWF_C.
DR Pfam; PF00007; Cys_knot; 1.
DR Pfam; PF00219; IGFBP; 1.
DR Pfam; PF00090; TSP_1; 1.
DR Pfam; PF00093; TSP_1; 1.
DR Pfam; PF00093; TSP_1; 1.
DR SMART; SM00041; CT; 1.
DR SMART; SM00121; IB; 1.
DR SMART; SM00209; TSPL; 1.
DR SMART; SM00214; VWF; 1.
DR PROSITE; PS01185; CTCK_1; 1.
DR PROSITE; PS01225; CTCK_2; 1.
DR PROSITE; PS00992; TSPL; 1.
DR PROSITE; PS01208; VWF; 1.
SO SEQUENCE 367 AA; 40702 MW; 3B7C0569EFAB5E96 CRC64;

Query Match 36.1%; Score 758.5; DB 11; Length 367;
Best Local Similarity 37.6%; Pred. No. 6e-64;
Matches 143; Conservative 55; Mismatches 135; Indels 47; Gaps 6;

3 SSTRFLAVAVTLHLTRALSTCPAACHCPLEAPKCAPGVGLVRDGGCCGCAKOLNE 62
26 SPTPTMTTPAPLETTTRPRECKKPCPCPSPPRCPLGVSLITDGCCECCICAQOL 85
63 DCSKTOPCDHTKLECNF--GASSTALKGICRAOSEGRPEYNSRIYONGESFQPNCKH 121
86 NCTEALICDPHRLGYCDYSGDRPRYALIGVC-AQVVGVCVLGVRNNGSQSPNCKYN 144
122 TCIDGAVGCIPLC--PQELSLPNLNGCPRPLVYKSGGCCCEWVDEDSIKDSLDDDDL 179
145 TCIDGAVGCTPLCLSPR---PLMCRQPRHVRVPEQCCQWVEDDADR----- 192
180 GLDASEVELTRNNELIAGKSSLKRLPYGTEPRVLFNPLAHGCKCIQVOTSMSCSK 239
193 -----RQTALDLDRFAASGAVGQRENCIATYSPMSPCST 228
240 SCGTGISTRVNTDNPFCRLVETRICVAPCGQPVYSSLKGRKSKTKKSPVPRFTYA 299
229 TCGIGISTRIISNNACWPEQESRLCNLRPCVDVIDLITLAKGKCLAVYQPEATNFTLA 288
300 GCSVYKYPKYGCGSVDRCCPTLOTTRYKMRFBEDGEMFSKNMIMOSCKNCPH 359
289 GCVSTRTYCPKYGVMDCNRCCLPYKSKTIDVSFCQDGLGFSROYLWLNACFCNLSCRN 348
360 PNEASFRLYSLFNDIHKFRD 379
349 PND-----IFADLESYPD 361

RESULT 11
099P0 PRELIMINARY; PRT; 367 AA.
ID 099P0;
AC 099P0;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE ELM1.
GN ELM1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=20487548; PubMed=11031104;
RA Sleeman M.A., Murtison J.G., Strachan L., Kumble K.D., Glenn M.P.,
RA McGraith A., Bickstaff P., Grierson A., Havukkala I., Tan P.,

RX MEDLINE=99061933; PubMed=9843955;
 RA Penica D., Swenson T.A., Welsh J.W., Roy M.A., Lawrence D.A., Lee J.,
 RA Brush J., Taneyhill L.A., Deuel B., Lew M., Matnabe C., Cohen R.L.,
 RA Welham M.F., Finley G.G., Quirke P., Goddard A.D., Hillan K.J.,
 RA Gurney A.L., Botstein D., Levine A.J.;
 RT "WSP genes are members of the connective tissue growth factor family
 RT that are up-regulated in wnt-1-transformed cells and aberrantly
 RT expressed in human colon tumors."
 RL Proc. Natl. Acad. Sci. U.S.A. 95:14717-14722(1998).
 DR EMBL: AF100781; AAC96323.1; -
 DR Gene: HGNC:12771; WISP3.
 DR InterPro: IPR006208; Cys_knot.
 DR InterPro: IPR006207; Cys_knot.C.
 DR InterPro: IPR000867; Insl_gro_fac-pr.
 DR InterPro: IPR000884; TSPL.
 DR Pfam: PF00007; Cys_knot.1.
 DR Pfam: PF00219; IGFBP.1.
 DR Pfam: PF00909; TSP.1; 1.
 DR SMART: SM0041; CT.1.
 DR SMART: SM00121; IB.1.
 DR SMART: SM00209; TSP1.1.
 DR PROSITE: PS01225; CTCK.2.1.
 DR PROSITE: PS00222; IGF_BINDING.1.
 DR PROSITE: PS50092; TSP1.1.
 DR PROSITE: PS50092; TSP1.1.
 SQ SEQUENCE 354 AA; 39292 MW; 67F48D0D5C2F5EE3 CRC64;

Query Match 29.6%; Score 622; DB 4; Length 354;

Best Local Similarity 34.4%; Pred. No. 5.9e-51;
Matches 118; Conservative 56; Mismatches 119; Indels 50; Gaps 8;

QY 26 CPACHCPLPAPKAPGVGLVNDGCGCCVCAKOLNEDSKTOPCDHTGLEGCFGASST 85
 DB 48 CHMPCPKPQKPCPCPGVSLVRDGGCCCKKCAKQPEICNEADLCDPHKGLCYDVSVD 107
 QY 86 ALK-GICRAQSGREPCVNSRTYONGESPONCKHCTCIDAVGCIPLCPQELSPNG 144
 DB 108 RIETGVC-AYLVAVGEEFQVHNMQVQPNPLFSCILVSGAIGCTP-----LFTPKLA 161
 QY 145 ---CPNRLVYKVGQCCCEWVDEDSIKSLDDDDLLGLDASEVELTRNNELIAGKGS 201
 DB 162 GSHCSGAK---GKKSDQSNCSLEPLDQLO-----ST 190
 QY 202 SLKRLPVECTEPRVLEPLHAHGOKCIYVTTSSQCSKSCGIGISTRVNDNPECKLYVE 261
 DB 191 SKYTMPAYRNPLI-----WKKKCLVQATKWTPTCSRTGKMGISNRVTNENSCMRKE 243
 QY 262 TRICEVRPGCGOPYSLK--KGGKCSKTEKKSPEPVFTYAGSSVKKYPRKYGCGVDNR 319
 DB 244 KRLCTIOPCDNSNLTKIKIPKGTCTPTOLSKAEKFEVSGSSTOSYKPTFCGICLDR 303
 QY 320 CCTPLQTRVKKMRFCEDEGEMFSKNYMTIQSCCKANYNCPHNE 362
 DB 304 CCIPNKSMTITQFDCPNESGFKWKMWTITSCVCQNCNCRPGD 346

RESULT 14

OBCIC8 PRELIMINARY; PRT; 251 AA.
 AC 08CIC8:
 DT 01-MAR-2003 (TREMBLrel. 23, Created)
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Similar to WNT1 inducible signaling pathway protein 2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Mammary gland;
 RA Strausberg R.;
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC032877; AAH32877.1; -

SQ SEQUENCE 251 AA; 27106 MW; 57520309A9069524 CRC64;

Query Match 24.7%; Score 518.5; DB 11; Length 251;
 Best Local Similarity 37.5%; Pred. No. 2.9e-41;
 Matches 100; Conservative 41; Mismatches 81; Indels 45; Gaps 5;

QY 15 LHLTLRLALS-----TCPACHCPLPAPKAPGVGLVNDGCGCCVCAKOLNED 63
 DB 6 LHLL--LAISFLICILISMVAYQALCPAPCACPMPPPQCPGVPLVLDGCGCCVCAKRLGSS 63
 QY 64 CSKTQPCDHTKGLGECNFGASSYALKICCAQSGRCCEVNSRTYONGESPONCKHCTQC 123
 DB 64 CCHLHVNCNPSQGVYCCPGAGPSGRCVCLFEEDDGSCEVNGRATYLDGETFKPCRYLRC 123
 QY 124 INGAVGCIPLCPQELSLPNIAGCPNPLVYKVGQCCCEWVDEDSIKSLDDDDLLGLDA 183
 DB 124 DDGFGTCPLCSDDVYLPMDCDPRPRIRIYPRGRCCEWVCDQAVMPAL----- 172
 QY 184 SEVELTRNNELIAGKSSILKRLPVGTEPRVLEPLHAHGOKCIYVTTSSQCSKSCGT 243
 DB 173 -----QPSAQ-----GHQLSLVTPASADG-PCPMWSTAMGPCSTTGL 211
 QY 244 GISTRYTNDNPECKLYKETRICVPRC 270
 DB 212 GIATRVSNQNRFCQLEIQRRLCLSRPC 238

RESULT 15

Q9HCS3 PRELIMINARY; PRT; 280 AA.

AC 09HCS3:
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE WISP-1 variant.
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Tanaka S., Sugimachi K.;
 RT "Human WISP-1 variant."
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB034725; BAB17849.1; -
 DR InterPro: IPR006208; Cys_knot.
 DR InterPro: IPR006207; Cys_knot.C.
 DR InterPro: IPR000867; Insl_gro_fac-pr.
 DR InterPro: IPR000884; TSPL.
 DR Pfam: PF00007; Cys_knot.1.
 DR Pfam: PF00219; IGFBP.1.
 DR Pfam: PF00909; TSP.1; 1.
 DR SMART: SM00041; CT.1.
 DR SMART: SM00121; IB.1.
 DR SMART: SM00209; TSP1.1.
 DR PROSITE: PS01185; CTCK.1.1.
 DR PROSITE: PS01225; CTCK.2.1.
 DR PROSITE: PS50092; TSP1.1.
 SQ SEQUENCE 280 AA; 30697 MW; 26B254DA060738E CRC64;

Query Match 23.8%; Score 499.5; DB 4; Length 280;

Best Local Similarity 28.6%; Pred. No. 2.1e-39;
Matches 108; Conservative 46; Mismatches 89; Indels 135; Gaps 8;

QY 9 LAVAVTILHLTLRLALST-----CPACHCPLPAPKAPGVGLVNDGCGCCVCAKOLN 61
 DB 25 LSPAPPTMDFTPAPELDTSSRPQFKWPCPCPPSPRCPGLSVLITDGECCCKMAQQLG 84
 QY 62 EDCSKTQPCDHTKGLGECNFGASSYALKICCAQSGRCCEVNSRTYONGESPONCKHCTQC 121
 DB 85 DICTEALICDPPHRLGLCYD-----SGDRP-----RYAIGV-----CAH-- 117
 QY 122 TCIDGAVGCIPLCPQELSLPNIAGCPNPLVYKVGQCCCEWVDEDSIKSLDDDDLLGL 181


```
Db      118  ----AVG----- 120
QY      182  DASFEVLTJRNNELAIIGKSSLKRLPVFGTEPRVLEPNLHAHGKCIIVOTISMSQCSKC 241
Db      121  ---EVEAMHRN----- 143
QY      242  GTGISTPVTNDNEPCRLVKEFRICQEVPRCGQPVYSSLKGGKCSKTKKSPPEVRFYAGC 301
Db      144  GLGVSTRISNVNACQPEQESRLCNLRPCVDYDHTILIKAGKKCLAVYQPEASMNFTLAGC 203
QY      302  SSVKKYRPKYCGSCVDGRCCCTPIQRTVKKMFRCEDGEMFSKNVMIQSCKNYNCPPN 361
Db      204  ISTRSYQPKYCGVCMNRCCIPYKSKTIDVSFQCPDGLGFSROYLWLNACFCNLSCRPN 263
QY      362  EASFRLXSLFNDIHKFRD 379
Db      264  D-----IFADLESYPD 274
```

Search completed: August 5, 2003, 14:06:50
Job time : 35.9132 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 5, 2003, 14:02:12 ; Search time 39.1026 Seconds
(without alignments)
1546.567 Million cell updates/sec

Title: US-09-495-448a-4
Perfect score: 2116
Sequence: 1 MSRRARALALVTLHLTR.....ANEAPFFYRLFNHFRD 381

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: A_Geneseq_19Jun03.*
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
23: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
24: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
25: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2116	100.0	381	AAW35730	Human cysteine ric
2	2116	100.0	381	AAE05921	Human cysteine-ric
3	2106	99.5	381	AAW35957	Human monocytic mat
4	2106	99.5	381	AAW30773	Human shear stress
5	2106	99.5	381	AAU79761	Human Cyr61 protei
6	2106	99.5	381	AAE18107	Human connective t
7	2106	99.5	381	ABB05438	Human Cyr61 protei
8	2106	99.5	455	AAW3987	Human cancer assoc
9	2106	99.1	381	ABG76937	Human protein, com

10	1971.5	93.2	374	23	ABB09202	HCGF CNN family pr
11	1971.5	93.2	375	20	AAW09919	Connective tissue
12	1971.5	93.2	375	20	AAW31620	Human CTGF-2. Hom
13	1971.5	93.2	375	23	AAE18108	Human alternative
14	1929	91.2	379	13	AAE25565	Beta-IG-M1. Mus m
15	1929	91.2	379	22	AAE05920	Mouse cysteine-ric
16	1929	91.2	379	23	ABB09201	Mouse cyr6 CNN fam
17	1695	80.1	375	22	AAE05939	Chicken CEP10 prot
18	1695	80.1	375	22	ABB09203	Chicken celo CNN f
19	953	45.0	347	20	AAI24379	Rat connective tis
20	952.5	45.0	348	13	AAW25566	Beta-IG-M2. Mus m
21	948.5	44.8	348	18	AAW35731	Murine fisp12. Mu
22	948.5	44.8	348	21	AAW44756	Mouse connective t
23	948.5	44.8	348	22	AAE05922	Mouse fibroblast s
24	948.5	44.8	348	23	ABB09205	Fisp-12 CNN family
25	948	44.8	347	18	AAW12694	Connective tissue
26	948	44.8	347	21	AAW33340	Amino acid sequenc
27	947	44.8	349	16	AAW79964	Connective tissue
28	947	44.8	349	18	AAW09089	Human connective t
29	947	44.8	349	18	AAW11302	Connective tissue
30	947	44.8	349	19	AAW62084	Human connective t
31	947	44.8	349	20	AAW18361	Human connective t
32	947	44.8	349	20	AAW81425	Connective tissue
33	947	44.8	349	21	AAW92939	Human connective t
34	947	44.8	349	21	AAW92940	Human connective t
35	947	44.8	349	21	AAW44755	Human connective t
36	947	44.8	349	22	AAE05923	Human connective t
37	947	44.8	349	22	AAW84598	Amino acid sequenc
38	947	44.8	349	22	AAW90791	Human shear stress
39	947	44.8	349	22	AAW60664	Human connective t
40	947	44.8	349	22	AAW88831	Human connective t
41	947	44.8	349	23	ABP68624	Human pancreatic c
42	938.5	44.4	348	23	ABB09204	Human ctyf CNN fam
43	869	40.2	349	21	AAW44754	Bovine connective
44	851.5	40.2	351	14	AAW31599	Chicken nov protei
45	851.5	40.2	351	22	AAE05940	Avian Nov protein.

ALIGNMENTS

RESULT 1
ID AAW35730
AAW35730 standard; Protein: 381 AA.
AC AAW35730;
DT 27-MAR-1998 (first entry)
XX
XX
DE Human cysteine rich protein 61 (Cyr61).
XX
KW Cysteine rich protein 61; Cyr61; human;
KW extracellular matrix signalling molecule; cell adhesion;
KW cell migration; cell proliferation; angiogenesis; chondrogenesis;
KW oncogenesis; haematostasis; wound healing; organ regeneration.
XX
OS Homo sapiens.
XX
PN W09733995-A2.
XX
PD 18-SEP-1997.
XX
PE 14-MAR-1997; 97WO-US04193.
XX
PR 15-MAR-1996; 96US-0013958.
XX
PA (MUNI-) MUNIN CORP.
XX
PI Lau LF.
XX
DR WPI: 1997-470875/43.
DR N-PSDB: AAT94699.
XX

PT Isolated and purified cysteine rich protein 61, Cyr61 - useful to
 PT modulate e.g. haemostasis, induce wound healing, promote organ
 PT regeneration etc
 PS Claim 2; Page 112-113; 133pp; English.
 XX
 XX This protein sequence comprises human cysteine rich protein 61
 CC (Cyr61), an extracellular matrix signalling molecule. Its amino
 CC acid sequence was deduced from a human placental cDNA clone (see
 CC AAF94699). Cyr61 polypeptides can be expressed in transformed or
 CC transfected host cells. Cyr61 can be used to modulate
 CC haemostasis, induce wound healing in a tissue, promote organ
 CC regeneration, improve tissue grafting or promote bone or prothesis
 CC implantation (claimed). It can also be used to screen for a
 CC modulator of angiogenesis, chondrogenesis, oncogenesis, cell
 CC adhesion, cell migration, cell proliferation, expand a population
 CC of undifferentiated haematopoietic stem cells in culture and to
 CC screen for a mitogen (claimed). Ex vivo methods for using
 CC mammalian extracellular matrix signalling molecules to prepare
 CC blood products are also provided.
 XX
 XX Sequence 381 AA;
 SQ
 Query Match 100.0%; Score 2116; DB 18; Length 381;
 Best Local Similarity 100.0%; Pred. No. 1.9e-162;
 Matches 381; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MSSRIARALALVYTLHLTRALSTCPACCHPLEAPKCAPGVGLVDRDGGCCKVCANOL 60
 DB 1 MSSRIARALALVYTLHLTRALSTCPACCHPLEAPKCAPGVGLVDRDGGCCKVCANOL 60
 QY 61 NEDCKSTQPCDHTKGLGECNFGASSTALKGICRAOSEGRPEYNSRIYQNGESFQPCNQHQ 120
 DB 61 NEDCKSTQPCDHTKGLGECNFGASSTALKGICRAOSEGRPEYNSRIYQNGESFQPCNQHQ 120
 QY 121 CTCIDAVAGCIRPLCPQELSLPNIGCPNPLVYKTGGCCCEWVCDDESIKPMEDDGLLG 180
 DB 121 CTCIDAVAGCIRPLCPQELSLPNIGCPNPLVYKTGGCCCEWVCDDESIKPMEDDGLLG 180
 QY 181 KELGFDASEVELTRNNELIANGKGRSLKRLPVFGMEPRILYNPLOGOKCIYOTTSMGQS 240
 DB 181 KELGFDASEVELTRNNELIANGKGRSLKRLPVFGMEPRILYNPLOGOKCIYOTTSMGQS 240
 QY 241 KTCGTGISTRYTNDNPECRVLKTRICEVRPGQPYVSSILKKKCKSKTKKSPVAFETY 300
 DB 241 KTCGTGISTRYTNDNPECRVLKTRICEVRPGQPYVSSILKKKCKSKTKKSPVAFETY 300
 QY 301 AGCLSVKVKYRPKCGSCVDRGCTPOLTRTVKMRFRCEDEFTSKNVMIOGCKNVCNP 360
 DB 301 AGCLSVKVKYRPKCGSCVDRGCTPOLTRTVKMRFRCEDEFTSKNVMIOGCKNVCNP 360
 QY 361 HANEAAFPFYRLFNDIHKFRD 381
 DB 361 HANEAAFPFYRLFNDIHKFRD 381
 RESULT 2
 AAE05921
 ID AAE05921 standard; Protein; 381 AA.
 XX
 AC AAE05921;
 XX
 DT 24-SEP-2001 (first entry)
 XX
 DE Human cysteine-rich protein (Cyr61).
 XX
 KM Cysteine-rich protein; Cyr61; extracellular matrix signalling molecule;
 KM fibroblast secreted protein; Fisp12; connective tissue growth factor;
 KM CTGF; ECM; cell adhesion; cell migration; fibroblast cell proliferation;
 KM angiogenesis; wound healing; integrin receptor; atherosclerosis; tumour;
 KM heart disease; fibrosis; gene therapy; human.
 XX
 OS Homo sapiens.

XX
 FH Key Location/Qualifiers
 FT Region 163..229
 FT /note="Cysteine free region"
 FT Domain 212..281
 FT /note="Domain III"
 FT Domain 282..381
 FT /note="Domain IV"
 XX
 PN W0200155210-A2.
 XX
 PD 02-AUG-2001.
 XX
 PF 31-JAN-2001; 2001WO-0503267.
 XX
 PR 31-JAN-2000; 2000US-0495448.
 PR 15-MAY-2000; 2000US-0204364.
 PR 06-OCT-2000; 2000US-0238705.
 XX
 XX (MUNN-) MUNN CORP.
 PA
 PI Lau LF, Yeung C, Greenspan JA;
 XX
 XX WPI; 2001-465561/50.
 DR N-PSDB; AAD11221.
 XX
 PT Novel human cysteine-rich protein 61 (Cyr61) fragment useful in methods
 PT for screening for modulators of cell adhesion, fibroblast cell
 PT proliferation, angiogenesis and cell migration
 XX
 PS Claim 30; Page 171-172; 186pp; English.
 XX
 XX The invention relates to extracellular matrix (ECM) signalling
 CC molecules involved in cellular response to growth factors. More
 CC particularly the invention is directed to cysteine-rich protein
 CC (Cyr61), and Cyr61-related proteins such as fibroblast secreted
 CC protein (Fisp12) and connective tissue growth factor (CTGF) and
 CC nucleic acid molecules encoding such proteins. The polypeptides
 CC of the invention are members of cysteine-rich secreted protein
 CC family. Human Cyr61 fragment is useful in methods for screening
 CC modulators of cell adhesion, cell migration, fibroblast cell
 CC proliferation, angiogenesis, wound healing and Cyr61-integrin
 CC receptor interaction. Modulator of Cyr61-integrin alphavbeta3
 CC interaction is used for the preparation of a medicament for the
 CC treatment of atherosclerosis, heart disease, tumour metastasis,
 CC fibrosis, tumour growth, disorders associated with inadequate
 CC angiogenesis; aberrant granulation tissue development; aberrant
 CC fibroblast growth and wounds. Polynucleotides of the invention
 CC are useful in gene therapy. The present sequence is human Cyr61
 CC protein.
 XX
 SQ Sequence 381 AA;
 Query Match 100.0%; Score 2116; DB 22; Length 381;
 Best Local Similarity 100.0%; Pred. No. 1.9e-162;
 Matches 381; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MSSRIARALALVYTLHLTRALSTCPACCHPLEAPKCAPGVGLVDRDGGCCKVCANOL 60
 DB 1 MSSRIARALALVYTLHLTRALSTCPACCHPLEAPKCAPGVGLVDRDGGCCKVCANOL 60
 QY 61 NEDCKSTQPCDHTKGLGECNFGASSTALKGICRAOSEGRPEYNSRIYQNGESFQPCNQHQ 120
 DB 61 NEDCKSTQPCDHTKGLGECNFGASSTALKGICRAOSEGRPEYNSRIYQNGESFQPCNQHQ 120
 QY 121 CTCIDAVAGCIRPLCPQELSLPNIGCPNPLVYKTGGCCCEWVCDDESIKPMEDDGLLG 180
 DB 121 CTCIDAVAGCIRPLCPQELSLPNIGCPNPLVYKTGGCCCEWVCDDESIKPMEDDGLLG 180
 QY 181 KELGFDASEVELTRNNELIANGKGRSLKRLPVFGMEPRILYNPLOGOKCIYOTTSMGQS 240
 DB 181 KELGFDASEVELTRNNELIANGKGRSLKRLPVFGMEPRILYNPLOGOKCIYOTTSMGQS 240

QY 241 KTCGTGISTRYTNDNPECRIVKETRICVRCGQPVYSLKGGKCKSKTKKSPERVPTTY 300
 DB 241 KTCGTGISTRYTNDNPECRIVKETRICVRCGQPVYSLKGGKCKSKTKKSPERVPTTY 300
 QY 301 AGCLSVKKYRPRKYGSCVDGRCCTPQLTRIVKMRFCEDGETFSKNVMMIOSCKKNYCP 360
 DB 301 AGCLSVKKYRPRKYGSCVDGRCCTPQLTRIVKMRFCEDGETFSKNVMMIOSCKKNYCP 360
 QY 361 HANEAFPEYRLFNDIHKFRD 381
 DB 361 HANEAFPEYRLFNDIHKFRD 381

RESULT 3

AAM35957
 ID AAM35957 standard; Protein: 381 AA.

AAW35957;

05-MAR-1998 (first entry)

Human monocyte mature differentiation factor.

Human; monocyte; mature; differentiation factor; MMD; macrophage;
 cancer; immune activator; tissue culture; infectious disease.

Homo sapiens.

JP09234079-A.

09-SEP-1997.

04-MAR-1996; 96JP-0075236.

04-MAR-1996; 96JP-0075236.

(TOYM) TOYOBO KK.

WPI: 1997-497320/46.

N-PSDB: AAT97142.

A monocyte mature differentiation factor - useful for the long term
 tissue culture of macrophage(s)

Claim 9; Page 12-13; 22pp; Japanese.

The present sequence represents a monocyte mature differentiation
 factor (MMD) which maintains the life of macrophages for long periods
 in liquid culture. MMD can be used as an anti-cancer agent, an immune
 activator and to treat infectious diseases.

Sequence 381 AA;

Query Match 99.5%; Score 2106; DB 18; Length 381;
 Best Local Similarity 99.5%; Pred. No. 1,2e-161;
 Matches 379; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSSRIARALAVTLHLTRLALSTCPAACHCPLEAPKACAGVGLVBDGCGCCCKVCAKOL 60
 DB 1 MSSRIARALAVTLHLTRLALSTCPAACHCPLEAPKACAGVGLVBDGCGCCCKVCAKOL 60
 QY 61 NEDSKTOPCDHRTKGLCNGFASSTALKGICRAOSEGRPEYNSRITONGESFQPNCKHQ 120
 DB 61 NEDSKTOPCDHRTKGLCNGFASSTALKGICRAOSEGRPEYNSRITONGESFQPNCKHQ 120
 QY 121 CTCIDGAVGCIPLCPQELSLPNLGCNPRIVKVTGQCCCEWVDCEDSIXKPMEDQDLG 180
 DB 121 CTCIDGAVGCIPLCPQELSLPNLGCNPRIVKVTGQCCCEWVDCEDSIXKPMEDQDLG 180
 QY 181 KKLGFDASEVELTRNNELIANGKRSILKRLPVFGMEBRILYNPLQGGKCIYQTTSMWGS 240
 DB 181 KKLGFDASEVELTRNNELIANGKRSILKRLPVFGMEBRILYNPLQGGKCIYQTTSMWGS 240

QY 241 KTCGTGISTRYTNDNPECRIVKETRICVRCGQPVYSLKGGKCKSKTKKSPERVPTTY 300
 DB 241 KTCGTGISTRYTNDNPECRIVKETRICVRCGQPVYSLKGGKCKSKTKKSPERVPTTY 300
 QY 301 AGCLSVKKYRPRKYGSCVDGRCCTPQLTRIVKMRFCEDGETFSKNVMMIOSCKKNYCP 360
 DB 301 AGCLSVKKYRPRKYGSCVDGRCCTPQLTRIVKMRFCEDGETFSKNVMMIOSCKKNYCP 360
 QY 361 HANEAFPEYRLFNDIHKFRD 381
 DB 361 HANEAFPEYRLFNDIHKFRD 381

RESULT 4

AAB90773
 ID AAB90773 standard; Protein: 381 AA.

AAB90773;

15-JUN-2001 (first entry)

Human shear stress-response protein SEQ ID NO: 46.

Human; shear stress-response protein; vascular disease;
 arteriosclerosis.

Homo sapiens.

WO200125427-A1.

12-APR-2001.

02-OCT-2000; 2000WO-JP06840.

01-OCT-1999; 99JP-0280976.

(KYOW) KYOWA HAKKO KOGYO KK.

(NOJI/) NOJIMA H.

Nojima H, Yoshisue H, Obayashi M, Ota T, Kawabata A, Sakurada K;
 Kuga T, Sekine S, Nakamura Y, Sugano S;

WPI: 2001-266308/27.

N-PSDB: AAH02896.

DNA sequences, proteins encoded by them and antibodies against them
 useful in diagnosis and treatment of vascular disease caused by
 arteriosclerosis -

Claim 60; Page 345-346; 678pp; Japanese.

The present invention provides the protein and coding sequences of a
 number of human shear stress response proteins. These are useful in the
 diagnosis, treatment and screening of vascular diseases caused by
 arteriosclerosis, including heart failure, post-PTCA restenosis and
 hypertension.

Sequence 381 AA;

Query Match 99.5%; Score 2106; DB 22; Length 381;
 Best Local Similarity 99.5%; Pred. No. 1,2e-161;
 Matches 379; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSSRIARALAVTLHLTRLALSTCPAACHCPLEAPKACAGVGLVBDGCGCCCKVCAKOL 60
 DB 1 MSSRIARALAVTLHLTRLALSTCPAACHCPLEAPKACAGVGLVBDGCGCCCKVCAKOL 60
 QY 61 NEDSKTOPCDHRTKGLCNGFASSTALKGICRAOSEGRPEYNSRITONGESFQPNCKHQ 120
 DB 61 NEDSKTOPCDHRTKGLCNGFASSTALKGICRAOSEGRPEYNSRITONGESFQPNCKHQ 120
 QY 121 CTCIDGAVGCIPLCPQELSLPNLGCNPRIVKVTGQCCCEWVDCEDSIXKPMEDQDLG 180
 DB 121 CTCIDGAVGCIPLCPQELSLPNLGCNPRIVKVTGQCCCEWVDCEDSIXKPMEDQDLG 180

Db 121 CTCIDGAVGCIPLCPQELSLPNIIGCPNRLVKTGQCCERWVCEDESIKDPMEDDGLLG 180
 Qy 181 KELGFDASEVELTNNELIIVGKGRSLKRLPVFGMEPRILYINPLOGKCIYVOTTSMSQCS 240
 Db 181 KELGFDASEVELTNNELIIVGKGRSLKRLPVFGMEPRILYINPLOGKCIYVOTTSMSQCS 240
 Qy 241 KTCGTGISTRTVNDNPECRIVKETRICEVRPCGQPVYSSILKKGKSKTKKSPPEVRYTY 300
 Db 241 KTCGTGISTRTVNDNPECRIVKETRICEVRPCGQPVYSSILKKGKSKTKKSPPEVRYTY 300
 Qy 301 AGCLSVKRYKPKYKCGSCVDGRCTPOLTRTVKMRFCEDGETFSKNVMMIOSCKCNYNCP 360
 Db 301 AGCLSVKRYKPKYKCGSCVDGRCTPOLTRTVKMRFCEDGETFSKNVMMIOSCKCNYNCP 360
 Qy 361 HANEAPFPYRLFNDIHKFRD 381
 Db 361 HANEAPFPYRLFNDIHKFRD 381

RESULT 5
 AAD79761
 ID AAD79761 standard; Protein; 381 AA.

AC AAD79761;
 XX 30-JUL-2002 (first entry)
 DT
 XX
 DE Human Cyr61 protein.
 XX

KW Human; uterine leiomyoma proliferation; uterine leiomyoma formation;
 KM Cyr61; cytosolic.
 XX

OS Homo sapiens.
 OS
 PN W0200226193-A2.
 PD 04-APR-2002.

PF 28-SEP-2001; 2001WO-US30783.
 XX
 PR 29-SEP-2000; 2000US-236887P.

XX (AMHP) AMERICAN HOME PROD CORP.

PI Zhang Z, Sampath D, Zhu Y, Wineker R;

DR WPI: 2002-383245/41.
 DR N-PSDB; ABK48899.

XX Preventing uterine leiomyoma formation or inhibiting proliferation of
 PT uterine leiomyoma in subject, comprises modulating or increasing the
 PT level of Cyr61 in leiomyoma tissue

XX Disclosure; Fig 6; 92pp; English.

XX The present invention relates to a method of inhibiting proliferation
 CC of uterine leiomyoma or preventing uterine leiomyoma formation. The
 CC method comprises increasing the level of Cyr61 in leiomyoma tissue.
 CC The invention also describes compounds and compositions that stimulate
 CC induction of the Cyr61 gene and compounds that increase Cyr61 activity.
 CC The compositions and the method of the invention are useful for
 CC preventing uterine leiomyoma formation or inhibiting proliferation of
 CC uterine leiomyoma in a subject. The method is particularly useful for
 CC treating or preventing uterine leiomyoma formation, or inhibiting
 CC proliferation of uterine leiomyoma in a subject. The present sequence
 CC represents human Cyr61.
 CC Note: The present sequence shown in Fig 6 is not shown in the correct
 CC sequence order in the figure. The start of the sequence is shown on
 CC page 8/10 of the figures and the rest of the sequence is shown on
 CC page 6/10 of the figures.

XX Sequence 381 AA;
 SO

Query Match 99.5%; Score 2106; DB 23; Length 381;
 Best Local Similarity 99.5%; Pred. No. 1, 2e-161;
 Matches 379; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MSSRIARALALVYTLHLTPALSTCPAACHCLEAPKCAPGVGLRDGGCCCAKOL 60
 Db 1 MSSRIARALALVYTLHLTPALSTCPAACHCLEAPKCAPGVGLRDGGCCCAKOL 60
 Qy 61 NECCSKTOPCDHTKGLCECNFGASTALKGICRAQSEGRPCENSRITYONGESFQPNCOHQ 120
 Db 61 NECCSKTOPCDHTKGLCECNFGASTALKGICRAQSEGRPCENSRITYONGESFQPNCKHQ 120
 Qy 121 CTCIDGAVGCIPLCPQELSLPNIIGCPNRLVKTGQCCERWVCEDESIKDPMEDDGLLG 180
 Db 121 CTCIDGAVGCIPLCPQELSLPNIIGCPNRLVKTGQCCERWVCEDESIKDPMEDDGLLG 180
 Qy 181 KELGFDASEVELTNNELIIVGKGRSLKRLPVFGMEPRILYINPLOGKCIYVOTTSMSQCS 240
 Db 181 KELGFDASEVELTNNELIIVGKGRSLKRLPVFGMEPRILYINPLOGKCIYVOTTSMSQCS 240
 Qy 241 KTCGTGISTRTVNDNPECRIVKETRICEVRPCGQPVYSSILKKGKSKTKKSPPEVRYTY 300
 Db 241 KTCGTGISTRTVNDNPECRIVKETRICEVRPCGQPVYSSILKKGKSKTKKSPPEVRYTY 300
 Qy 301 AGCLSVKRYKPKYKCGSCVDGRCTPOLTRTVKMRFCEDGETFSKNVMMIOSCKCNYNCP 360
 Db 301 AGCLSVKRYKPKYKCGSCVDGRCTPOLTRTVKMRFCEDGETFSKNVMMIOSCKCNYNCP 360
 Qy 361 HANEAPFPYRLFNDIHKFRD 381
 Db 361 HANEAPFPYRLFNDIHKFRD 381

RESULT 6
 AAE18107
 ID AAE18107 standard; Protein; 381 AA.

AC AAE18107;

DT 07-MAY-2002 (first entry)
 XX

DE Human connective tissue growth factor-2 (CTGF-2).

XX Human; angiogenesis; connective tissue growth factor-2; CTGF-2; tumour;
 XX ischaemia; restenosis; tissue repair; wound healing; congenital defect;
 KW cardiovascular disease; atherosclerosis; heart failure; angina; trauma;
 KW burns; osteoporosis; periodontal disease; liver failure; tranquilizer;
 KW vulnery; cosmetic plastic surgery; vasotropic; hepatotropic; ulcer;
 KW gene therapy.
 KM
 OS Homo sapiens.
 OS

XX Location/Qualifiers

FT Region 35..41 "Immunogenic epitope"
 FT /note="Immunogenic epitope"
 FT Region 47..51 "Immunogenic epitope"
 FT /note="Immunogenic epitope"
 FT Region 59..75 "Immunogenic epitope"
 FT /note="Immunogenic epitope"
 FT Region 91..119 "Immunogenic epitope"
 FT /note="Immunogenic epitope"
 FT Region 145..150 "Immunogenic epitope"
 FT /note="Immunogenic epitope"
 FT Region 164..176 "Immunogenic epitope"
 FT /note="Immunogenic epitope"
 FT Region 202..208 "Immunogenic epitope"
 FT /note="Immunogenic epitope"
 FT Region 223..228 "Immunogenic epitope"
 FT /note="Immunogenic epitope"
 FT Region 239..244 "Immunogenic epitope"
 FT /note="Immunogenic epitope"
 FT Region 250..257 "Immunogenic epitope"
 FT /note="Immunogenic epitope"

FT Region 279..296
 FT /note= "Immunogenic epitope"
 FT Region 307..314
 FT /note= "Immunogenic epitope"
 FT Region 318..323
 FT /note= "Immunogenic epitope"
 FT Region 337..343
 FT /note= "Immunogenic epitope"
 XX
 PN WO200204480-A2.
 XX
 PD 17-JAN-2002.
 XX
 PE 11-JUL-2001; 2001WO-US21799.
 XX
 PR 11-JUL-2000; 2000US-217402P.
 PR 18-MAY-2001; 2001US-291642P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 PA (TRGE) TRANSGENE SA.
 XX
 PI Li H, Adams MD, Calenda V, Fatacciolli V;
 DR WPI; 2002-171698/22.
 DR N-PDB; AAD29095.
 XX
 PT Stimulating angiogenesis in a mammal preferably human having ischemia
 PT or restenosis or is treated for limb revascularization, by
 PT administering connective tissue growth factor-2 polypeptide or
 PT polynucleotide
 XX
 PS Example 1; Fig 1; 131pp; English.
 XX
 CC The present invention relates to a method for stimulating angiogenesis in
 CC a mammal. The method comprises administering a polynucleotide encoding
 CC connective tissue growth factor-2 (CTGF-2) or an active fragment or its
 CC derivative. The method is useful for stimulating angiogenesis in a mammal
 CC preferably human having ischemia or restenosis or is treated for limb
 CC revascularization which is leg or arm. The invention is useful for
 CC inhibiting tumor growth, where angiogenesis is utilised for enhancing
 CC the repair of connective and support tissue, promoting the attachment,
 CC fixation and stabilisation of tissue implants and enhancing wound
 CC healing, hence is useful for treating cardiovascular disease e.g.
 CC atherosclerosis, reperfusion injury such as heart failure, angina,
 CC ischemia; and is also used to differentiate, proliferate and attract
 CC cells leading to regeneration of tissues which is utilised to repair
 CC replace or protect tissue damaged by congenital defects, trauma (burns,
 CC ulcer, etc), age, disease (e.g. osteoporosis, periodontal disease,
 CC liver failure), surgery including cosmetic plastic surgery. The present
 CC sequence is human CTGF-2. CTGF-2 gene is useful in gene therapy.
 CC
 SO Sequence 381 AA;
 Query Match 99.5%; Score 2106; DB 23; Length 381;
 Best Local Similarity 99.5%; Pred. No. 1.2e-161;
 Matches 379; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 241 KTCGTGISTRVNDNPECRLVETRICVPRCCGPVYSLLKGGKCKSKTKSPRPVFTY 300
 DB 241 KTCGTGISTRVNDNPECRLVETRICVPRCCGPVYSLLKGGKCKSKTKSPRPVFTY 300
 QY 301 AGCLSVKRYKRYKCGSCVDCRCCTPOLTRVKNMFRCEDEFTSKNNM10SCKNYNC 360
 DB 301 AGCLSVKRYKRYKCGSCVDCRCCTPOLTRVKNMFRCEDEFTSKNNM10SCKNYNC 360
 QY 361 HANEAFFPYRLFNDRHFRD 381
 DB 361 HANEAFFPYRLFNDRHFRD 381
 RESULT 7
 ABB05438
 ID ABB05438 standard; Protein: 381 AA.
 XX
 AC ABB05438;
 XX
 DT 15-APR-2002 (first entry)
 XX
 DE Human Cyr61 protein SEQ ID NO:2.
 XX
 KW Human; Cyr61; breast cancer; sex steroid receptor; cytosolic; promoter;
 KW sex steroid response element; cysteine rich heparin-binding protein;
 KW cell proliferation; heparin binding epidermal growth factor;
 KW epidermal growth factor; basic fibroblastic growth factor.
 XX
 OS Homo sapiens.
 XX
 PN WO200198359-A2.
 PD 27-DEC-2001.
 XX
 PE 21-JUN-2001; 2001WO-US19823.
 XX
 PR 21-JUN-2000; 2000US-213182P.
 PR 16-MAY-2001; 2001US-291510P.
 XX
 PA (AMHP) AMERICAN HOME PROD CORP.
 XX
 PI Sampath D, Zhang Z, Winneker R;
 DR WPI; 2002-147796/19.
 DR N-PDB; ABA93127, ABA93130.
 XX
 PT Regulation of Cyr61 expression and activity for preventing and
 PT inhibiting breast cancer comprises use of a Cyr61 neutralizing
 PT antibody, an anti-sense oligonucleotide and an antibody which
 XX
 PS Claim 6; Fig 1; 86pp; English.
 XX
 CC The present invention describes a method for the prevention or inhibition
 CC of breast cancer cell proliferation. The method comprises administration
 CC of a compound that inhibits the interaction of a sex steroid receptor
 CC with a sex steroid response element of the Cyr61 (cysteine rich heparin-
 CC binding protein) promoter. Cyr61 has cytosolic activity. An antibody (1)
 CC which neutralises Cyr61 can be used to prevent or inhibit breast cancer
 CC cell proliferation by blocking sex steroid induced and growth factor
 CC induced synthesis of Cyr61 DNA, where the growth factor is epidermal,
 CC heparin binding epidermal or basic fibroblastic growth factor. (1) can be
 CC used to diagnose or stage breast cancer where the level of Cyr61 in a
 CC positive/suspect breast cancer cell is compared to the level in a
 CC normal cell; an increase in the level of Cyr61 compared to the level in
 CC normal tissue indicates the presence of breast cancer. The level of
 CC Cyr61 being determined by exposing the tissues to (1), and an increase
 CC in the level of bound antibody by the suspect/positive cell as compared
 CC to the normal tissue indicates the presence of breast cancer. The present
 CC sequence represents the human Cyr61 protein, which is used in the
 CC exemplification of the present invention.
 CC
 SO Sequence 381 AA;

Query Match 99.5%; Score 2106; DB 23; Length 381;
 Best Local Similarity 99.5%; Pred. No. 1.2e-161;
 Matches 379; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSSRIARALAVVTLHLRLALSTCPAACHCPLPAPKAPGVGLVRDGGCCCKVCARQL 60
 DB 1 MSSRIARALAVVTLHLRLALSTCPAACHCPLPAPKAPGVGLVRDGGCCCKVCARQL 60
 QY 61 NEDCSKTQPCDHTKLECNFASSTALKGICRAOSEGRPCENSRITYNGESFQPCNQHQ 120
 DB 61 NEDCSKTQPCDHTKLECNFASSTALKGICRAOSEGRPCENSRITYNGESFQPCNQHQ 120
 QY 121 CTCIDGAVGICPLCPQELSLPMLGCPNRLVYVGTGCCCEWYCDSDSIKDPMEDDGLLG 180
 DB 121 CTCIDGAVGICPLCPQELSLPMLGCPNRLVYVGTGCCCEWYCDSDSIKDPMEDDGLLG 180
 QY 181 KELGFDASEVELTRNNELIANGKGRSLKRLPVFGMEPRILYNPLDGGKCIYOTTSWSQCS 240
 DB 181 KELGFDASEVELTRNNELIANGKGRSLKRLPVFGMEPRILYNPLDGGKCIYOTTSWSQCS 240
 QY 241 KTCGTGISTRYTNDNPECRIVKETRICEVRPCGQPVYSSLKKGKCSKTKSPPEVRYTY 300
 DB 241 KTCGTGISTRYTNDNPECRIVKETRICEVRPCGQPVYSSLKKGKCSKTKSPPEVRYTY 300
 QY 301 AGCLSVKKRYKRYKCGSCVDGRCTPOLTRTYVMRRFCEDGETFSKNVMYIOSCKCNYNCP 360
 DB 301 AGCLSVKKRYKRYKCGSCVDGRCTPOLTRTYVMRRFCEDGETFSKNVMYIOSCKCNYNCP 360
 QY 361 HANEAAFPFYRLFNDIHKFRD 381
 DB 361 HANEAAFPFYRLFNDIHKFRD 381

RESULT 8
 AAB43987
 ID AAB43987 standard; Protein: 455 AA.
 AC AAB43987;
 DT 08-FEB-2001 (first entry)
 DE Human cancer associated protein sequence SEQ ID NO:1432.
 XX

Human; cancer associated gene; cancer antigen; detection; cancer;
 diagnosis; cytostatic; proliferative; vulnerrary; immunomodulator;
 antidiabetic; antiaschmatic; antirheumatic; antitubercitic; antiviral;
 dermatological; antitubercitic; antidiabetic; antitubercitic; antitubercitic;
 vasotrophic; antiprotective; thrombolytic; coagulant; neutropenic;
 immune disorder; haematopoietic cell disorder; autoimmune disorder;
 allergic reaction; graft versus host disease; organ rejection;
 haemostatic; thrombolytic; cardiovascular disorder; infection;
 neurological disease; drug screening.

OS Homo sapiens.
 XX
 PN WO20005350-A1.
 XX
 PD 21-SEP-2000.
 XX
 PF 08-MAR-2000; 2000WO-US05882.
 XX
 PR 12-MAR-1999; 99US-0124270.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Ruben SM;
 XX
 XX WPI: 2000-587533/55.
 DR N-PSDB: AAC78196.
 XX
 PT Novel isolated nucleic acids comprising sequences encoding peptides
 useful for treating or diagnosing e.g. cancer -

XX Claim 11; Page 2116-2118; 2352pp; English.
 PS AAC7607 to AAC78448 encode the human cancer associated proteins given
 XX in AAB43987 to AAB44239. The proteins can have activities based on the
 CC tissues and cells the genes are expressed in. Example of activities
 CC include: cytostatic; proliferative; vulnerrary; immunomodulator;
 CC antidiabetic; antiaschmatic; antirheumatic; antitubercitic;
 CC antiinflammatory; antitubercitic; antidiabetic; antitubercitic;
 CC dermatological; neuroprotective; cardiant; thrombolytic; coagulant;
 CC neutropenic; vasotrophic; antiprotective and antitubercitic. The
 CC polynucleotides and polypeptides can be used for preventing, treating or
 CC ameliorating medical conditions and diagnosing pathological conditions.
 CC Polynucleotides, polypeptides, antibodies, agonists and antagonists from
 CC the present invention may be used to treat immune disorders by activating
 CC or inhibiting the proliferation, differentiation or mobilisation of
 CC immune cells, to treat disorders of haematopoietic cells, autoimmune
 CC disorders, allergic reactions, graft versus host disease and organ
 CC rejection, modulate haemostatic or thrombolytic activity, modulate
 CC inflammation, cancers, cardiovascular disorders, neurological disease and
 CC bacterial or viral infections. The peptides, nucleotides, antibodies,
 CC agonists and antagonists may be also be used in drug screens. AAC78449 to
 CC AAC78457 and AAB44240 represent sequences used in the exemplification of
 CC the present invention.

Sequence 455 AA:
 QY 1 MSSRIARALAVVTLHLRLALSTCPAACHCPLPAPKAPGVGLVRDGGCCCKVCARQL 60
 DB 75 MSSRIARALAVVTLHLRLALSTCPAACHCPLPAPKAPGVGLVRDGGCCCKVCARQL 134
 QY 61 NEDCSKTQPCDHTKLECNFASSTALKGICRAOSEGRPCENSRITYNGESFQPCNQHQ 120
 DB 135 NEDCSKTQPCDHTKLECNFASSTALKGICRAOSEGRPCENSRITYNGESFQPCNQHQ 194
 QY 121 CTCIDGAVGICPLCPQELSLPMLGCPNRLVYVGTGCCCEWYCDSDSIKDPMEDDGLLG 180
 DB 195 CTCIDGAVGICPLCPQELSLPMLGCPNRLVYVGTGCCCEWYCDSDSIKDPMEDDGLLG 254
 QY 181 KELGFDASEVELTRNNELIANGKGRSLKRLPVFGMEPRILYNPLDGGKCIYOTTSWSQCS 240
 DB 255 KELGFDASEVELTRNNELIANGKGRSLKRLPVFGMEPRILYNPLDGGKCIYOTTSWSQCS 314
 QY 241 KTCGTGISTRYTNDNPECRIVKETRICEVRPCGQPVYSSLKKGKCSKTKSPPEVRYTY 300
 DB 315 KTCGTGISTRYTNDNPECRIVKETRICEVRPCGQPVYSSLKKGKCSKTKSPPEVRYTY 374
 QY 301 AGCLSVKKRYKRYKCGSCVDGRCTPOLTRTYVMRRFCEDGETFSKNVMYIOSCKCNYNCP 360
 DB 375 AGCLSVKKRYKRYKCGSCVDGRCTPOLTRTYVMRRFCEDGETFSKNVMYIOSCKCNYNCP 434
 QY 361 HANEAAFPFYRLFNDIHKFRD 381
 DB 435 HANEAAFPFYRLFNDIHKFRD 455

RESULT 9
 ABG76937
 ID ABG76937 standard; Protein: 381 AA.
 AC ABG76937;
 DT 05-NOV-2002 (first entry)
 DE Human protein, comprising CYR61, designated SECL.
 XX
 XX
 XX
 XX
 DE Human protein, comprising CYR61, designated SECL.
 KW Human; SEC; NOV; immunosuppressive; hepatotropic;
 KW antiinflammatory; angiogenic-associated disorder; diagnostic;
 KW gene therapy; developmental disorder; immune disease;

KW signal transduction pathway disorder; metabolic disorder;
 KW feeding disorder; obesity; wasting disorder; neurodegenerative disorder;
 KW Alzheimer's disease; Parkinson's disease; behavioural disorder; allergy;
 KW asthma; atherosclerosis; cardiomyopathy; angina pectoris;
 KW autoimmune disease; retinal disease; cirrhosis; diabetes;
 KW infectious disease; human immunodeficiency virus; HIV; cancer;
 KW hypertension; hypotension; multiple sclerosis; urinary retention;
 KW osteoporosis; Crohn's disease; ulcer; neurological disorder; anxiety;
 KW hemophilia; cirrhosis; immunogen; vaccine.
 OS Homo sapiens.
 PN WO20025705-A2.
 XX 18-JUL-2002.
 PD 11-JAN-2002; 2002WO-US00609.
 PF 11-JAN-2001; 2001US-261013P.
 PR 11-JAN-2001; 2001US-261014P.
 PR 11-JAN-2001; 2001US-261018P.
 PR 11-JAN-2001; 2001US-261026P.
 PR 11-JAN-2001; 2001US-261029P.
 PR 17-AUG-2001; 2001US-313170P.
 PR 10-SEP-2001; 2001US-318410P.
 XX (CURA-) CURAGEN CORP.
 PA Mezes PS, Rastell L, Hermann JL, MacDougall JR, Zhong H;
 PI Casman SJ, Boldo F, Shmukets RA, Gorman L, Crasta OR, Mysore KK;
 PI Folkerts O, Martin GB, Eisen A, Spaderna SK, Vermet CAM, Bergh C;
 PI Spletter KA, DiPippo VA, Zernusen BD, Peyman JA, Ellemann K;
 PI Stone DJ, Grosse WM, Alsobrook JP, Lepley DM, Rieger DK;
 PI Burgess CE, Edinger S.
 XX WPI: 2002-590675/63.
 DR N-PSDB; ABS59522.
 XX Human SECX/NOVX polypeptide useful for diagnosing, preventing or
 PT treating disorders associated with aberrant expression or activity of
 PT SECX/NOVX nucleic acids and proteins e.g., diabetes
 XX Example 3; Page 9; 443pp; English.
 PS The invention discloses the isolated human polypeptides, and
 CC polynucleotides encoding them, that have been designated SECX and NOVX.
 CC The polypeptides can be used for treating, or delaying, the onset of an
 CC antigenic-associated disorder or treating a pathological state in a
 CC subject, preferably a mammal. They can also be used in determining the
 CC presence of, or predisposition to, a disease associated with altered
 CC levels of the polypeptides and polynucleotides of any one of the 12
 CC sequences (SEC1-12), for raising antibodies, for identifying an agent
 CC that binds to, or that modulates the expression or activity of the
 CC polypeptide, for treating or preventing a NOVX-associated disorder
 CC (NOV1-8) and as a pharmaceutical composition comprising the polypeptide,
 CC polynucleotide or the antibody. The polypeptides and polynucleotides are
 CC useful in diagnostic applications where their amounts are assessed, or
 CC for the manufacture of a medicament (e.g. gene therapy) for treating or
 CC preventing disorders or syndromes such as developmental disorders, immune
 CC diseases, signal transduction pathway disorders, metabolic disorders,
 CC feeding disorders (including obesity), wasting disorders,
 CC neurodegenerative disorders (including Alzheimer's disease and
 CC Parkinson's disease), behavioural disorders, allergies, asthma,
 CC atherosclerosis, cardiomyopathy, angina pectoris, autoimmune diseases,
 CC retinal disease, cirrhosis, diabetes, infectious disease (bacterial,
 CC fungal, protozoal and viral e.g. human immunodeficiency virus, HIV),
 CC cancer (e.g. prostate cancer), hypertension, hypotension, multiple
 CC sclerosis, urinary retention, osteoporosis, Crohn's disease, ulcers,
 CC neurological disorders (e.g. anxiety), haemophilia or cirrhosis. They
 CC may also be used as immunogens to produce antibodies specific for the
 CC invention, and as vaccines. Further, they are useful for screening
 CC potential agonist and antagonist compounds. The sequences presented in
 CC ABG76937-ABG76956 are the human SEC1-12 and NOV1-8 proteins.

XX Sequence 381 AA;
 SO Query Match 99.1%; Score 2098; DB 23; Length 381;
 Best Local Similarity 99.0%; Pred. No. 5,4e-161;
 Matches 377; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 1 MSSRIARALALVYVLLHTRIALSTCPAACHCPLEAKCAPGVGVBDGCGCKVCAKOL 60
 DB 1 MSSRIARALALVYVLLHTRIALSTCPAACHCPLEAKCAPGVGVBDGCGCKVCAKOL 60
 QY 61 NEDCSKTQPCDHTKGLCECNFGASSTALKGICRAOSEGRPCENSRITYONGSEFOPNCOHQ 120
 DB 61 NEDCSKTQPCDHTKGLCECNFGASSTALKGICRAOSEGRPCENSRITYONGSEFOPNCKHQ 120
 QY 121 CTCIDGAVGCIPLCPQELSLPNLGCPNRLVKYVKGCCCEWVCDSDSIKDPMDQDGLG 180
 DB 121 CTCIDGAVGCIPLCPQELSLPNLGCPNRLVKYVKGCCCEWVCDSDSIKDPMDQDGLG 180
 QY 181 KELGFDAEVELTFRNNELIANGKRSILKRLPVFPMEDRIILYNPLOGOKIQTWSWCS 240
 DB 181 KELGFDAEVELTFRNNELIANGKRSILKRLPVFPMEDRIILYNPLOGOKIQTWSWCS 240
 QY 241 KTCGTGISTRTYNDNPECRILVKEIRICEVPRGQPVYSLSLKKGKSKTKKSPVPRFTY 300
 DB 241 KTCGTGISTRTYNDNPECRILVKEIRICEVPRGQPVYSLSLKKGKSKTKKSPVPRFTY 300
 QY 301 AGCLSVKRYRKYKGCSCVDGRCTPOLTRVYKMFRCDEGETFSKNVMIOSCKNANCP 360
 DB 301 AGCLSVKRYRKYKGCSCVDGRCTPOLTRVYKMFRCDEGETFSKNVMIOSCKNANCP 360
 QY 361 HANEAPFPYRLFNDIHKFRD 381
 DB 361 HANEAPFPYRLFNDIHKFRD 381
 RESULT 10
 ABB09202
 ID ABB09202 standard; Protein: 374 AA.
 XX ABB09202;
 AC 08-JUL-2002 (first entry)
 DT HCGF CNM family protein sequence SEQ ID NO:12.
 XX Human; small CCN-like growth factor; SCGF; vulnery; osteopathic;
 KW gene therapy; muscle wasting disease; osteoporosis; wound healing;
 KW tissue regeneration; angiogenesis.
 OS Unidentified.
 XX US2002049304-A1.
 PN 25-APR-2002.
 PD 14-MAY-2001; 2001US-0853625.
 PF 06-JUN-1995; 95US-0468847.
 PR 01-APR-1998; 98US-0053587.
 XX (HAST/) HASTINGS G A.
 PA (ADAM/) ADAMS M D.
 XX Hastings GA, Adams MD.
 DR WPI: 2002-382150/41.
 XX Novel isolated polynucleotide sequence encoding a human small CCN-like
 PT growth factor, useful for treating muscle wasting disease, and
 PT osteoporosis -
 XX Disclosure; Fig 2A-D; 33pp; English.

XX The present invention describes human small CCN-like growth factor
 CC (SCGF). SCGF has vulnereary and osteopathic activities, and can be used
 CC in gene therapy. The SCGF polypeptides and polynucleotides can be used
 CC for treating muscle wasting diseases, and osteoporosis, and to stimulate
 CC wound healing and tissue regeneration, to promote angiogenesis and to
 CC stimulate proliferation of vascular smooth muscle and endothelial cell
 CC production. The present sequence represents a CCN family protein which
 CC is given in comparison with the human SCGF in the exemplification of the
 CC present invention.

XX Sequence 374 AA:

Query Match 93.2%; Score 1971.5; DB 23; Length 374;

Best Local Similarity 95.7%; Pred. No. 8.5e-151;

Matches 358; Conservative 2; Mismatches 13; Indels 1; Gaps 1;

QY 1 MSSRIARALAVYTLHLRLALSTCPACHCPLKAPGVLVBDGCGCCYCAKOL 60
 DB 1 MSSRIARALAVYTLHLRLALSTCPACHCPLKAPGVLVBDGCGCCYCAKOL 60
 QY 61 NEDCKTOPCDHTKGLKCNFGASSTALKGICRAOSBGRCEYNSRIYONGESFOPNCQHO 120
 DB 61 NEDCKTOPCDHTKGLKCNFGASSTALKGICRAOSBGRCEYNSRIYONGESFOPNCQHO 120
 QY 121 CTCIDGAVG-CIPLCPOELSLPMLGCPNRLVKTGOCCEWVCDSDSKDPMEDODGLI 179
 DB 121 CTCIDGAVG-CIPLCPOELSLPMLGCPNRLVKTGOCCEWVCDSDSKDPMEDODGLI 180
 QY 180 GKELGFDASEVELTRNNELIANGKRSRLKLPFGMEPRILYNPLOGOKCIYOTTWSQOC 239
 DB 181 GKELGFDASEVELTRNNELIANGKRSRLKLPFGMEPRILYNPLOGOKCIYOTTWSQOC 240
 QY 240 SKTCGTGISTRYTNDNPECRLVETRICVRCPCGVYSSLLKGGKCKSTKSPPEVPRFT 299
 DB 241 SKTCGTGISTRYTNDNPECRLVETRICVRCPCGVYSSLLKGGKCKSTKSPPEVPRFT 300
 QY 300 YAGCLSVKKYRPRKYGSCVDGRCTPQTLRTYKMFRCDEGTFESKNVMIOGSKCNVNC 359
 DB 301 YAGCLSVKKYRPRKYGSCVDGRCTPQTLRTYKMFRCDEGTFESKNVMIOGSKCNVNC 360
 QY 360 PHANEAAPFYRLF 373
 DB 361 PHANEAAPFYRLF 374

RESULT 11

AA90919 standard; Protein; 375 AA.

XX AA90919;

XX 25-MAR-2003 (updated)

XX 25-JUN-1996 (first entry)

XX Connective tissue growth factor-2.

XX CTGF-2; connective tissue growth factor-2; secreted protein;

XX cartilaginous growth; skeletal; embryo; cell growth; morphogenesis;

XX insulin-like growth factor; fibroblast growth factor; Crys1.

XX Homo sapiens.

XX Key Location/Qualifiers

XX Peptide 1...24

XX Protein /label= signal_peptide

XX WO9601896-A1.

XX 25-JAN-1996.

PF 12-JUL-1994; 94WO-0507736.
 XX
 PR 12-JUL-1994; 94WO-0507736.
 XX
 PA (HUMA-) HUMAN GENOME SCI. 18C.
 XX
 PI Adams MD, Li H;
 XX
 DR WPI: 1996-097626/10.
 DR N-PSDB; AAT12653.
 XX
 PT Connective tissue growth factor-2 and DNA encoding it - useful to
 PT enhance the repair of connective and support tissue, and to enhance
 PT wound healing
 PS
 PS Claim 1: Fig 1A-C; 46pp; English.

CC Connective tissue growth factor-2 (CTGF-2) is encoded by a cDNA
 CC (AAT12653) isolated from a human foetal lung cDNA library. The CTGF
 CC polypeptides are structurally and functionally related to a family
 CC of growth factors which include IGF (insulin-like growth factor),
 CC PDGF (platelet-derived growth factor), and FGF (fibroblast growth
 CC factor). CTGF-2 exhibits 89 percent identity and 93 percent similarity
 CC to Crys1. Crys1 is a growth factor-inducible immediate early gene
 CC initially identified in serum-stimulated mouse fibroblasts. It encodes
 CC a member of an emerging family of secreted proteins which are also a
 CC group of cysteine-rich proteins. This group of GFs are important for
 CC normal growth, differentiation, morphogenesis of the cartilaginous
 CC skeleton of an embryo and cell growth.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 XX
 SQ Sequence 375 AA:

Query Match 93.2%; Score 1971.5; DB 17; Length 375;

Best Local Similarity 95.7%; Pred. No. 8.5e-151;

Matches 358; Conservative 2; Mismatches 13; Indels 1; Gaps 1;

QY 1 MSSRIARALAVYTLHLRLALSTCPACHCPLKAPGVLVBDGCGCCYCAKOL 60
 DB 1 MSSRIARALAVYTLHLRLALSTCPACHCPLKAPGVLVBDGCGCCYCAKOL 60
 QY 61 NEDCKTOPCDHTKGLKCNFGASSTALKGICRAOSBGRCEYNSRIYONGESFOPNCQHO 120
 DB 61 NEDCKTOPCDHTKGLKCNFGASSTALKGICRAOSBGRCEYNSRIYONGESFOPNCQHO 120
 QY 121 CTCIDGAVG-CIPLCPOELSLPMLGCPNRLVKTGOCCEWVCDSDSKDPMEDODGLI 179
 DB 121 CTCIDGAVG-CIPLCPOELSLPMLGCPNRLVKTGOCCEWVCDSDSKDPMEDODGLI 180
 QY 180 GKELGFDASEVELTRNNELIANGKRSRLKLPFGMEPRILYNPLOGOKCIYOTTWSQOC 239
 DB 181 GKELGFDASEVELTRNNELIANGKRSRLKLPFGMEPRILYNPLOGOKCIYOTTWSQOC 240
 QY 240 SKTCGTGISTRYTNDNPECRLVETRICVRCPCGVYSSLLKGGKCKSTKSPPEVPRFT 299
 DB 241 SKTCGTGISTRYTNDNPECRLVETRICVRCPCGVYSSLLKGGKCKSTKSPPEVPRFT 300
 QY 300 YAGCLSVKKYRPRKYGSCVDGRCTPQTLRTYKMFRCDEGTFESKNVMIOGSKCNVNC 359
 DB 301 YAGCLSVKKYRPRKYGSCVDGRCTPQTLRTYKMFRCDEGTFESKNVMIOGSKCNVNC 360
 QY 360 PHANEAAPFYRLF 373
 DB 361 PHANEAAPFYRLF 374

RESULT 12

AA931620 standard; Protein; 375 AA.

XX AA931620;

XX 02-NOV-1999 (first entry)

XX Human CTGF-2.
 DE Connective tissue growth factor-2; CTGF-2; wound healing; bone disorder;
 XX skin disorder; acne; burn; UV damage; stabilisation; tissue implant.
 KW Homo sapiens.
 XX
 XX
 FT Key Location/Qualifiers
 FT Peptide 1..24
 FT Protein /label= Signal_peptide
 FT MISC-difference 25..375
 FT /note= "Cys encoded by ICT"
 FT
 PN US945300-A.
 XX
 PD 31-AUG-1999.
 XX
 PF 02-JUN-1995; 95US-0459101.
 XX
 PR 02-JUN-1995; 95US-0459101.
 PR 12-JUL-1994; 94WO-US07736.
 XX
 PA (ADAM/) ADAMS M D.
 PA (LITH/) LI H.
 XX
 PI Adams MD, LI H;
 XX
 DR WPI: 1999-508171/42.
 DR N-PSDB: AA211720.
 XX
 PT Polynucleotides encoding growth factor polypeptides useful for
 PT enhancing the repair of connective tissue and support tissue
 XX
 PS Claim 1; Fig 1; 20pp; English.
 XX
 CC This sequence represents human connective tissue growth factor-2
 CC (CTGF-2). CTGF-2 cDNA was isolated from a cDNA library derived from
 CC human foetal lung. In one instance, the cDNA was cloned into a
 CC baculovirus expression vector, having first been amplified and modified
 CC via PCR using primers AA211721 and AA211722. In another instance, the
 CC cDNA was cloned into a COS cell expression vector, with prior
 CC amplification and modification using PCR primers AA211723 and AA211724.
 CC CTGF-2 is structurally and functionally related to a family of growth
 CC factors which include IGF (insulin-like growth factor), PDGF
 CC (platelet-derived growth factor) and FGF (fibroblast growth factor). This
 CC emerging family of cysteine-rich secreted proteins are important for
 CC normal growth, differentiation, morphogenesis of the cartilaginous
 CC skeleton of an embryo and cell growth. Their functions also include wound
 CC healing, tissue repair, implant fixation and stimulating increased bone
 CC mass. CTGF-2 may be used to enhance the repair of connective tissue and
 CC support tissue and can therefore treat skin disorders e.g., acne, aging,
 CC UV damage or burns. CTGF-2 can be used to promote the attachment,
 CC fixation and stabilisation of tissue implants inserted during
 CC reconstructive surgery, and can be used to enhance the healing of
 CC external wounds. It can be used in the treatment of injured or depleted
 CC bone as it promotes the growth of connective tissue, bone and cementum
 CC and stimulates protein and collagen synthesis.
 CC
 SQ Sequence 375 AA;
 XX
 Query Match 93.2%; Score 1971.5; DB 20; Length 375;
 Best Local Similarity 95.7%; Pred. No. 8.5e-151;
 Matches 358; Conservative 2; Mismatches 13; Indels 1; Gaps 1;
 QY 1 MSSRIARALAVVTLHLTRIALSTCPACHCPLEAPKAPGVGLVBDGCGCCVKCAKOL 60
 DB 1 MSSRIARALAVVTLHLTRIALSTCPACHCPLEAPKAPGVGLVBDGCGCCVKCAKOL 60
 QY 61 NEQSKRQPDHRRKGLECNAGSSTALKGICRAOSEPPEYNSRIRYONGBEFOPNCOHQ 120
 DB 61 NEQSKRQPDHRRKGLECNAGSSTALKGICRAOSEPPEYNSRIRYONGBEFOPNCKHQ 120

QY 121 CTCIDGAVG-CIPLCPQELSLPNTGCPNPRVLVKTGOCCEHWGDEDSIKDPMEDQGLL 179
 DB 121 CTCIGMRGACIPLCPQELSLPNTGCPNPRVLVKTGOCCEHWGDEDSIKDPMEDQGLL 180
 QY 180 GKELGPDASEVELTRNNELIYVAGKRSILKRLPVEGMERILYNNPLOGKCIYQTTSMSC 239
 DB 181 GKELGPDASEVELTRNNELIYVAGKRSILKRLPVEGMERILYNNPLOGKCIYQTTSMSC 240
 QY 240 SKTCGTGTSTRTVTDNPECRILVETRICENAPCCQPYSSILKKKKSKTKSPPEVFT 299
 DB 241 SKTCGTGTSTRTVTDNPECRILVETRICENAPCCQPYSSILKKKKSKTKSPPEVFT 300
 QY 300 YAGCLSVKKRYPKRCGSCVDRCCTPOLTRIVKKRFRGDEFTSKNNVMQSKCNYNC 359
 DB 301 YAGCLSVKKRYPKRCGSCVDRCCTPOLTRIVKKRFRGDEFTSKNNVMQSKCNYNC 360
 QY 360 PHANEAAFPFYRLF 373
 DB 361 PHANEAAFPFYRLF 374
 RESULT 13
 ID AAE18108 standard; Protein: 375 AA.
 AC AAE18108;
 DT 07-MAY-2002 (first entry)
 DE Human alternative connective tissue growth factor-2 (CTGF-2).
 XX
 KW Human: angiogenesis; connective tissue growth factor-2; CTGF-2; tumour;
 KW ischaemia; restenosis; tissue repair; wound healing; congenital defect;
 KW cardiovascular disease; atherosclerosis; heart failure; angina; trauma;
 KW burns; osteoporosis; periodontal disease; liver failure; tranquilizer;
 KW vulnery; cosmetic plastic surgery; vasotropic; hepatotropic; ulcer;
 KW gene therapy.
 XX
 OS Homo sapiens.
 PN WO200204480-A2.
 PD 17-JAN-2002.
 PF 11-JUL-2001; 2001WO-US21799.
 PR 11-JUL-2000; 2000US-217402P.
 PR 18-MAY-2001; 2001US-291642P.
 PA (HUMA-) HUMAN GENOME SCI INC.
 PA (TRGE) TRANSGENE SA.
 PI LI H, Adams MD, Calenda V, Fataccioli V;
 DR WPI: 2002-171698/22.
 DR N-PSDB: AAD29099.
 PT Stimulating angiogenesis in a mammal preferably human having ischemia
 PT or restenosis or is treated for limb revascularization, by
 PT administering connective tissue growth factor-2 polypeptide or
 PT polynucleotide
 PS Disclosure; Fig 11; 13pp; English.
 CC The present invention relates to a method for stimulating angiogenesis in
 CC a mammal. The method comprises administering a polynucleotide encoding
 CC connective tissue growth factor-2 (CTGF-2) or an active fragment or its
 CC derivative. The method is useful for stimulating angiogenesis in a mammal
 CC preferably human having ischemia or restenosis or is treated for limb
 CC revascularisation which is leg or arm. The invention is useful for
 CC inhibiting tumour growth, where angiogenesis is utilised for enhancing
 CC the repair of connective and support tissue, promoting the attachment,

CC fixation and stabilisation of tissue implants and enhancing wound
 CC healing, hence is useful for treating cardiovascular disease e.g.
 CC atherosclerosis, reperfusion injury such as heart failure, angina,
 CC ischaemia; and is also used to differentiate, proliferate and attract
 CC cells leading to regeneration of tissues which is utilised to repair
 CC replace or protect tissue damaged by congenital defects, trauma (burns,
 CC ulcer, etc), age, disease (e.g. osteoporosis, periodontal disease,
 CC liver failure), surgery including cosmetic plastic surgery. The present
 CC sequence is human alternative CTGF-2. CTGF-2 gene is useful in gene
 CC therapy.

XX Sequence 375 AA:

Query Match 93.2%; Score 1971.5; DB 23; Length 375;

Best Local Similarity 95.7%; Pred. No. 8.5e-151;
 Matches 358; Conservative 2; Mismatches 13; Indels 1; Gaps 1;

QY 1 MSSRIARALAVYTLHLTRALSTCPACHCPLKAPKAPGVGLVDDGGCCCKVAKOL 60
 DB 1 MSSRIARALAVYTLHLTRALSTCPACHCPLKAPKAPGVGLVDDGGCCCKVAKOL 60
 QY 61 NEDCSKTQPCDHTKGLKCNFAGASTALKGICRAOSGRCPEYNSRIYONGESFOPNCOHQ 120
 DB 61 NEDCSKTQPCDHTKGLKCNFAGASTALKGICRAOSGRCPEYNSRIYONGESFOPNCOHQ 120
 QY 121 CTCIDGAVG-CIPCLPQELSLPNIGCPNRLVVTGQCCBEMWCDEDSIKDPMEDDGLL 179
 DB 121 CTCIDGAVG-CIPCLPQELSLPNIGCPNRLVVTGQCCBEMWCDEDSIKDPMEDDGLL 180
 QY 180 GRELGPASVEYELTRNNELIYAVGKSLKRLPVFGMEPRLLYPLGOKCIYOTTSWOC 239
 DB 181 GRELGPASVEYELTRNNELIYAVGKSLKRLPVFGMEPRLLYPLGOKCIYOTTSWOC 240
 QY 240 SKTCGTGISTRTVNDNPECKRLVETRIKCEVRPCGOPYSSILKGGKCSKTKKPEPRFT 299
 DB 241 SKTCGTGISTRTVNDNPECKRLVETRIKCEVRPCGOPYSSILKGGKCSKTKKPEPRFT 300
 QY 300 YAGCLSVKRYRKYCGSCVDGRCTPOLRTVYKMFRCDEDETFSKNVMYIOSCKNYNC 359
 DB 301 YAGCLSVKRYRKYCGSCVDGRCTPOLRTVYKMFRCDEDETFSKNVMYIOSCKNYNC 360
 QY 360 PHANEAFPPYRLF 373
 DB 361 PHANEAFPPYRLF 374

RESULT 14

AAR25565 standard; Protein: 379 AA.

XX AAR25565;
 XX 25-MAR-2003 (updated)
 DT 18-JAN-1993 (first entry)
 XX Beta-IG-M1.
 XX Transforming growth factor beta; induced; CEF-10; v-src; chicken;
 KM embryo; fibroblasts; TGF-beta.
 XX Mus musculus.
 PN EP495674-A2.
 XX 22-JUL-1992.
 XX 17-JAN-1992; 92EP-0300429.
 XX 18-JAN-1991; 91US-0642991.
 PR 10-JAN-1992; 92US-0816270.
 XX (BRIM) BRISTOL-MYERS SQUIBB CO.

PI Brunner AM, Chinn J, Neubauer MG, Purchio AF;
 XX WPL; 1992-243508/30.
 DR N-PSDB; AAQ26421.
 XX TGF-beta induced gene family - encodes proteins involved in
 PT growth and differentiation effects of TGF-beta-1
 PS Claim 2; Fig 1; 35pp; English.

CC The protein sequence was deduced from the DNA sequence obtd. by
 CC screening a cDNA library made from AKR-2B mouse cells induced with
 CC TGF-beta1 and cyclohexamide with two probes from untreated AKR-2B
 CC mRNA and AKR-2B mRNA from cells treated with cyclohexamide and TGF-
 CC beta1. The proteins encoded by hybridising colonies (beta-IG-M1 and
 CC beta-IG-M2) contain 38 Cys residues and are induced by TGF-beta1.
 CC Beta-IG-M1 displays 80 percent homology to the CEF-10 protein
 CC induced by v-src in chicken embryo fibroblasts and is identical
 CC to the protein encoded by cyf61, an immediate early response gene
 CC induced in quiescent BALB 3T3 cells by serum treatment. Residues
 CC 49-56 of beta-IG-M1 conform to the GCGCCXXC motif reported in the
 CC amino half of insulin-like growth factor (IGF) binding proteins.
 CC The C-terminal Cys rich region of beta-IG-M1, -M2 and CEF-10 contain
 CC an amino acid sequence with strong homology to a motif found near the
 CC C-terminal of the malarial circumsporozoite (CS) protein, which is
 CC highly conserved among all species of malarial parasites sequenced
 CC to date (designated region II). This motif is also found in
 CC other proteins which have cell adhesive properties that mediate
 CC cell-cell and cell-extracellular matrix interactions, such as
 CC perlecan, thrombospondin, and TRAP. The proteins encoded by
 CC TGF-beta induced genes are likely to be involved in mediation of
 CC the biological effects of TGF-beta relating to cell growth and
 CC differentiation. See also AAR25566.
 CC (Updated on 25-MAR-2003 to correct PN field.)

XX Sequence 379 AA:

Query Match 91.2%; Score 1929; DB 13; Length 379;

Best Local Similarity 90.9%; Pred. No. 2.3e-147;
 Matches 348; Conservative 10; Mismatches 19; Indels 6; Gaps 2;

QY 1 MSSRIARALAVYTLHLTRALSTCPACHCPLKAPKAPGVGLVDDGGCCCKVAKOL 60
 DB 1 MSSRIARALAVYTLHLTRALSTCPACHCPLKAPKAPGVGLVDDGGCCCKVAKOL 60
 QY 61 NEDCSKTQPCDHTKGLKCNFAGASTALKGICRAOSGRCPEYNSRIYONGESFOPNCOHQ 120
 DB 61 NEDCSKTQPCDHTKGLKCNFAGASTALKGICRAOSGRCPEYNSRIYONGESFOPNCOHQ 120
 QY 121 CTCIDGAVG-CIPCLPQELSLPNIGCPNRLVVTGQCCBEMWCDEDSIKDPMEDDGLL 180
 DB 121 CTCIDGAVG-CIPCLPQELSLPNIGCPNRLVVTGQCCBEMWCDEDSIKDPMEDDGLL 180
 QY 181 KEGLPASVEYELTRNNELIYAVGKSLKRLPVFGMEPRLLYPLGOKCIYOTTSWOC 238
 DB 179 --LGLDASEVELTRNNELIYAVGKSLKRLPVFGMEPRLLYPLGOKCIYOTTSWOC 236
 QY 239 CSKTCGTGISTRTVNDNPECKRLVETRIKCEVRPCGOPYSSILKGGKCSKTKKPEPRFT 298
 DB 237 CSKTCGTGISTRTVNDNPECKRLVETRIKCEVRPCGOPYSSILKGGKCSKTKKPEPRFT 296
 QY 299 TYAGCLSVKRYRKYCGSCVDGRCTPOLRTVYKMFRCDEDETFSKNVMYIOSCKNYNC 358
 DB 297 TYAGCLSVKRYRKYCGSCVDGRCTPOLRTVYKMFRCDEDETFSKNVMYIOSCKNYNC 356
 QY 359 PHANEAFPPYRLF 381
 DB 357 PHANEAFPPYRLF 379

RESULT 15

AAE05920 standard; Protein: 379 AA.

XX AA05920;
 AC 24-SEP-2001 (first entry)
 DT Mouse cysteine-rich protein (Cyr61).
 DE
 XX Cysteine-rich protein; Cyr61; extracellular matrix signalling molecule;
 KM fibroblast secreted protein; fisp12; connective tissue growth factor;
 KM CTGF; ECM: cell adhesion; cell migration; fibroblast cell proliferation;
 KM angiogenesis; wound healing; integrin receptor; atherosclerosis; tumour;
 KM heart disease; fibrosis; gene therapy; mouse.
 XX Mus musculus.
 OS
 XX
 FH Key Location/Qualifiers
 FT Region 164..226
 FT /note="Cysteine free region"
 FT 224..240
 FT Domain /note="Domain IIF"
 XX
 XX WO200155210-A2.
 PN
 XX
 PD 02-AUG-2001.
 XX
 PF 31-JAN-2001; 2001WO-US03267.
 XX
 PR 31-JAN-2000; 2000US-0495448.
 PR 15-MAY-2000; 2000US-0204364.
 PR 06-OCT-2000; 2000US-0238705.
 XX
 PA (MUNI-) MUNIN CORP.
 XX
 PI Lau LF, Yeung C, Greenspan JA;
 XX
 DR WPI: 2001-465561/50.
 DR N-PSDB: AAD11220.
 XX
 PT Novel human cysteine-rich protein 61 (Cyr61) fragment useful in methods
 PT for screening for modulators of cell adhesion, fibroblast cell
 PT proliferation, angiogenesis and cell migration
 XX
 PS Example 1; Fig 1; 186pp; English.
 XX
 CC The invention relates to extracellular matrix (ECM) signalling
 CC molecules involved in cellular response to growth factors. More
 CC particularly the invention is directed to cysteine-rich protein
 CC (Cyr61), and Cyr61-related proteins such as fibroblast secreted
 CC protein (fisp12) and connective tissue growth factor (CTGF) and
 CC nucleic acid molecules encoding such proteins. The polypeptides
 CC of the invention are members of cysteine-rich secreted protein
 CC family. Human Cyr61 fragment is useful in methods for screening
 CC modulators of cell adhesion, cell migration, fibroblast cell
 CC proliferation, angiogenesis, wound healing and Cyr61-integrin
 CC receptor interaction. Modulator of Cyr61-integrin alphavbeta3
 CC interaction is used for the preparation of a medicament for the
 CC treatment of atherosclerosis, heart disease, tumour metastasis,
 CC fibrosis, tumour growth, disorders associated with inadequate
 CC angiogenesis; aberrant granulation tissue development; aberrant
 CC fibroblast growth and wounds. Polynucleotides of the invention
 CC are useful in gene therapy. The present sequence is mouse Cyr61
 CC protein.
 XX
 SO Sequence 379 AA;

Query Match 91.2%; Score 1929; DB 22; Length 379;
 Best local similarity 90.9%; Pred. No. 2.3e-147;
 Matches 348; Conservative 10; Mismatches 19; Indels 6; Gaps 2;

QY 1 MSSRIARALAVTLHLTRALSTCPRAHCPLKAPGVLVRDGGCCCKVCAKOL 60
 DB 1 MSSSTFRLAVAVTLHLTRALSTCPRAHCPLKAPGVLVRDGGCCCKVCAKOL 60

QY 61 NEDCSKTQPCDHTKGLGECNFGASSTALNGICRAQSEGRCEYNSRTYONGESFOPNCQO 120
 DB 61 NEDCSKTQPCDHTKGLGECNFGASSTALNGICRAQSEGRCEYNSRTYONGESFOPNCQO 120
 QY 121 CTCIDGAVGCIPLCQBELSLPMLGCPNRLVAVTGOCCCEWYCDDESDIKDPHEDDGLG 180
 DB 121 CTCIDGAVGCIPLCQBELSLPMLGCPNRLVAVTGOCCCEWYCDDESDIKDPHEDDGLG 180
 QY 181 KETGPDASEVELTRNNELIANGKRSIKRLPYEGMEPRILYNPL--QGOKCIYQTSMSQ 238
 DB 179 --LGIDASEVELTRNNELIANGKRSIKRLPYEGMEPRILYNPL--QGOKCIYQTSMSQ 236
 QY 239 CSKTGCTGISTRTVNDNPECRVLYKETRICEVPRCGQPYSSLSKGGKCSKTKKSPDYRF 298
 DB 237 CSKSGCTGISTRTVNDNPECRVLYKETRICEVPRCGQPYSSLSKGGKCSKTKKSPDYRF 296
 QY 299 TYAGCLSVKKYRPKYCGSCVDGRCTPOLTRTKMRFRCEDETFESKVMMIQSCKNYN 358
 DB 297 TYAGCSSVKYRPKYCGSCVDGRCTPOLTRTKMRFRCEDETFESKVMMIQSCKNYN 356
 QY 359 CPHANEAPFPYRLFNDIHKRPD 381
 DB 357 CPHNEASFRLYSLFNDIHKRPD 379

Search completed: August 5, 2003, 14:05:12
 Job time : 40.1026 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 5, 2003, 14:03:52 ; Search time 14.0368 Seconds
(without alignments)
1148.438 Million cell updates/sec

Title: US-09-495-448a-4

Perfect score: 2116

Sequence: 1 MSSRIARALALVLTILHLTR.....ANEAAPFYRLFNDRFRD 381

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PTOUS_COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2116	100.0	381	US-09-142-569-4	Sequence 4, Appl1
2	2106	99.5	381	US-09-348-815-2	Sequence 2, Appl1
3	1971.5	93.2	374	US-08-468-847B-12	Sequence 12, Appl1
4	1964.5	92.8	375	US-08-459-101A-2	Sequence 2, Appl1
5	1929	91.2	379	US-08-468-847B-11	Sequence 11, Appl1
6	1929	91.2	379	US-09-142-569-2	Sequence 2, Appl1
7	1695	80.1	375	US-08-468-847B-13	Sequence 13, Appl1
8	953	45.0	347	US-09-582-337-2	Sequence 2, Appl1
9	952.5	45.0	348	US-09-292-036-3	Sequence 3, Appl1
10	948.5	44.8	348	US-08-468-847B-15	Sequence 15, Appl1
11	948.5	44.8	347	US-09-142-569-6	Sequence 6, Appl1
12	948.5	44.8	347	US-09-187-478-2	Sequence 2, Appl1
13	948	44.8	347	US-09-292-036-2	Sequence 2, Appl1
14	947	44.8	349	US-08-167-628-2	Sequence 2, Appl1
15	947	44.8	349	US-08-386-680-2	Sequence 2, Appl1
16	947	44.8	349	US-08-459-717-2	Sequence 2, Appl1
17	947	44.8	349	US-08-712-302-2	Sequence 2, Appl1
18	947	44.8	349	US-08-880-031-2	Sequence 2, Appl1
19	947	44.8	349	US-09-054-358-2	Sequence 2, Appl1
20	947	44.8	349	US-09-097-179-2	Sequence 2, Appl1
21	947	44.8	349	US-09-054-274-2	Sequence 2, Appl1
22	947	44.8	349	US-09-080-715-2	Sequence 2, Appl1
23	947	44.8	349	US-09-056-704-2	Sequence 2, Appl1
24	947	44.8	349	US-09-292-036-4	Sequence 2, Appl1
25	947	44.8	349	US-09-253-316-26	Sequence 26, Appl1
26	947	44.8	349	US-09-142-569-8	Sequence 8, Appl1
27	947	44.8	349	US-09-461-688-2	Sequence 2, Appl1

28	947	44.8	349	PCR-US96-08140-2	Sequence 2, Appl1
29	938.5	44.4	348	US-08-468-847B-14	Sequence 14, Appl1
30	851.5	40.2	351	US-08-468-847B-16	Sequence 16, Appl1
31	827.5	39.1	357	US-08-468-847B-17	Sequence 17, Appl1
32	827.5	39.1	357	US-09-253-316-25	Sequence 25, Appl1
33	767	36.2	367	US-09-182-145-4	Sequence 4, Appl1
34	767	36.2	367	US-09-182-145-8	Sequence 8, Appl1
35	766	36.2	367	US-09-182-145-7	Sequence 7, Appl1
36	766	36.2	367	US-09-182-145-22	Sequence 22, Appl1
37	760.5	35.9	345	US-09-182-145-3	Sequence 3, Appl1
38	760.5	35.9	345	US-09-182-145-6	Sequence 6, Appl1
39	759.5	35.9	345	US-09-182-145-5	Sequence 5, Appl1
40	759.5	35.9	345	US-09-182-145-21	Sequence 21, Appl1
41	758	35.8	367	US-09-182-145-12	Sequence 12, Appl1
42	754.5	35.7	345	US-09-182-145-11	Sequence 11, Appl1
43	613	29.0	339	US-09-182-145-36	Sequence 36, Appl1
44	613	29.0	354	US-09-182-145-37	Sequence 37, Appl1
45	613	29.0	354	US-09-253-316-2	Sequence 2, Appl1

ALIGNMENTS

```
RESULT 1
US-09-142-569-4
; Sequence 4, Application US/09142569
; Patent No. 6413735
GENERAL INFORMATION:
APPLICANT: Lau, Lester F.
TITLE OF INVENTION: Extracellular Matrix Signalling Molecules
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESS: Marshall, O'Noole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/142,569
FILING DATE: 02-Apr-1999
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Clough, David W.
REGISTRATION NUMBER: 36,107
REFERENCE/DOCKET NUMBER: 28758/33766
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6500
TELEFAX: 312/474-0448
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 381 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: misc.feature
OTHER INFORMATION: "Human Cyt61 amino acid sequence"
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-142-569-4
Query Match 100.0%; Score 2116; DB 4; Length 381;
Best Local Similarity 100.0%; Pred. No. 7.8e-177;
Matches 381; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 MSSRIARALALVLTILHLTRALSTCPAACHCPLFAPKCAVGILVRDGCRCVCAKOL 60
|||||
```

Db 1 MSSRIARALALVYTLHLTRLALSTCPAACHCPLKAPGVLVDDGCCCKVCAKOL 60
QY 61 NEDCSKTOPCDHRTKGLGECNFGASSTALKGICRAOSGRPCPEVNSRIYONGESPQPNCHQ 120
Db 61 NEDCSKTOPCDHRTKGLGECNFGASSTALKGICRAOSGRPCPEVNSRIYONGESPQPNCHQ 120
QY 121 CTCIDGAVGCIPLCPDELSPNLGCPNRLVKTGOCCEBWVDEDSIKDPMEDODGLLG 180
Db 121 CTCIDGAVGCIPLCPDELSPNLGCPNRLVKTGOCCEBWVDEDSIKDPMEDODGLLG 180
QY 181 KELGPDASEVELTRNNELLAVGKRSILKRLPVFGMEPRILYNPLQGOKCIYQTTSSQCS 240
Db 181 KELGPDASEVELTRNNELLAVGKRSILKRLPVFGMEPRILYNPLQGOKCIYQTTSSQCS 240
QY 241 KTCGTGISTRTVNDNPECRLVKEETRICCEVRPCGOPYSSILKKGKCSKTKKSPVRFY 300
Db 241 KTCGTGISTRTVNDNPECRLVKEETRICCEVRPCGOPYSSILKKGKCSKTKKSPVRFY 300
QY 301 AGCLSVKRYKPKYKCGSCVDGRCTPOLTRTVKMRFCEDGETFSKNVMMIOSCKYVNC 360
Db 301 AGCLSVKRYKPKYKCGSCVDGRCTPOLTRTVKMRFCEDGETFSKNVMMIOSCKYVNC 360
QY 361 HANEAAPEYRLFNDIHKFRD 381
Db 361 HANEAAPEYRLFNDIHKFRD 381

RESULT 2

US-09-348-815-2
Sequence 2, Application US/09348815
Patent No. 6534630
GENERAL INFORMATION:
APPLICANT: LI, HAODONG
ADAMS, MARK D
TITLE OF INVENTION: CONNECTIVE TISSUE GROWTH FACTOR-2
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: HUMAN GENOME SCIENCES, INC.
STREET: 9410 KEY WEST AVENUE
CITY: ROCKVILLE
STATE: MD
COUNTRY: US
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/348,815
FILING DATE: 08-Jul-1999
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: JONATHAN L. KLEIN
REGISTRATION NUMBER: 41,119
REFERENCE/DOCKET NUMBER: PFI26PID1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 301-309-8504
TELEFAX: 301-309-8439
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 381 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-348-815-2

Query Match 99.5%; Score 2106; DB 4; Length 381;
Best local Similarity 99.5%; Pred. No. 5,8e-176;
Matches 379; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 MSSRIARALALVYTLHLTRLALSTCPAACHCPLKAPGVLVDDGCCCKVCAKOL 60

Db 1 MSSRIARALALVYTLHLTRLALSTCPAACHCPLKAPGVLVDDGCCCKVCAKOL 60
QY 61 NEDCSKTOPCDHRTKGLGECNFGASSTALKGICRAOSGRPCPEVNSRIYONGESPQPNCHQ 120
Db 61 NEDCSKTOPCDHRTKGLGECNFGASSTALKGICRAOSGRPCPEVNSRIYONGESPQPNCHQ 120
QY 121 CTCIDGAVGCIPLCPDELSPNLGCPNRLVKTGOCCEBWVDEDSIKDPMEDODGLLG 180
Db 121 CTCIDGAVGCIPLCPDELSPNLGCPNRLVKTGOCCEBWVDEDSIKDPMEDODGLLG 180
QY 181 KELGPDASEVELTRNNELLAVGKRSILKRLPVFGMEPRILYNPLQGOKCIYQTTSSQCS 240
Db 181 KELGPDASEVELTRNNELLAVGKRSILKRLPVFGMEPRILYNPLQGOKCIYQTTSSQCS 240
QY 241 KTCGTGISTRTVNDNPECRLVKEETRICCEVRPCGOPYSSILKKGKCSKTKKSPVRFY 300
Db 241 KTCGTGISTRTVNDNPECRLVKEETRICCEVRPCGOPYSSILKKGKCSKTKKSPVRFY 300
QY 301 AGCLSVKRYKPKYKCGSCVDGRCTPOLTRTVKMRFCEDGETFSKNVMMIOSCKYVNC 360
Db 301 AGCLSVKRYKPKYKCGSCVDGRCTPOLTRTVKMRFCEDGETFSKNVMMIOSCKYVNC 360
QY 361 HANEAAPEYRLFNDIHKFRD 381
Db 361 HANEAAPEYRLFNDIHKFRD 381

RESULT 3

US-08-468-847B-12
Sequence 12, Application US/08468847B
Patent No. 5780263
GENERAL INFORMATION:
APPLICANT: Hastings, Gregg A. and Adams, Mark D.
TITLE OF INVENTION: Human CCN-like Growth Factor
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: CABELLA, BYRNE, BAIN, GILFILLIAN,
ADDRESSEE: CECCHI, STEWART & OLSTEIN
STREET: 6 BECKER FARM ROAD
CITY: ROSELAND
STATE: NEW JERSEY
COUNTRY: USA
ZIP: 07068
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 INCH DISKETTE
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: WORD PERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/468,847B
FILING DATE: 6 June 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: MULLINS, J.G.
REGISTRATION NUMBER: 33,073
REFERENCE/DOCKET NUMBER: 325800-442
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1744
TELEFAX: 201-994-1700
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 374 AMINO ACIDS
TYPE: AMINO ACID
STRANDEDNESS:
TOPOLOGY: LINEAR
MOLECULE TYPE: PROTEIN
US-08-468-847B-12

Query Match 93.2%; Score 1971.5; DB 1; Length 374;

Best Local Similarity 95.7%; Pred. No. 3.1e-164;
Matches 358; Conservative 2; Mismatches 13; Indels 1; Gaps 1;

```

QY 1 MSSRRALAVLTLHLTRALSTCPACACPLEAPRCAPGVGLVRGCGCCCKVCAROL 60
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 MSSRRALAVLTLHLTRVGLSTCPADCHCPLAPRCAPGVGLVRGCGCCCKVCAROL 60
QY 61 NEDSKTOPCDHTKGLKLECNFGASSSTALGICRAOSEGRPCENSRITYNGESFQPCNOHQ 120
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 NEDCKTOPCDHTKGLKLECNFGASSSTALGICRAOSEGRPCENSRITYNGESFQPCNOHQ 120
QY 121 CTCIDGANG-CIPCLPOELSLPNLGCNPNRLVKTGQCCCEWVCDEDSIKDPMEDQDGL 179
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 121 CTCIGWRGACIPLCPQELSLPNLGCNPNRLVKTGQCCCEWVCDEDSIKDPMEDQDGL 180
QY 180 GKELGFDSAEVELTRNNELIANGKRSILKRLPVFGMEPRILYNPLQGGKCIYQTSWSQC 239
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 181 GKELGFDSAEVELTRNNELIANGKRSILKRLPVFGMEPRILYNPLQGGKCIYQTSWSQC 240
QY 240 SKTCGTGISTVTNDNPECRVLKETRICEVRPCGQPVYSSILKKGKSKTKKSPPEVRF 299
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 241 SKTCGTGISTVTNDNPECRVLKETRICEVRPCGQPVYSSILKKGKSKTKKSPPEVRF 300
QY 300 YAGCLSVKRYPRKYGSCVDGRCTPOLTRTVKMFRCDEDETSKNVMYIOSCKCNYNC 359
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 301 YAGCLSVKRYPRKYGSCVDGRCTPOLTRTVKMFRCDEDETSKNVMYIOSCKCNYNC 360
QY 360 PHANEAAPFYRLF 373
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 361 PHANEAAPFYRLF 374

```

RESULT 4

US-08-459-101a-2
Sequence 2, Application US/08459101a

Patent No. 5945300

GENERAL INFORMATION:

APPLICANT: LI, ET AL.

TITLE OF INVENTION: Connective Tissue Growth Factor-2

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,

ADDRESSEE: CECCHI, STEWART & OLSTEIN

STREET: 6 BECKER FARM ROAD

CITY: ROSELAND

STATE: NEW JERSEY

COUNTRY: USA

ZIP: 07068

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 INCH DISKETTE

COMPUTER: IBM PS/2

OPERATING SYSTEM: MS-DOS

SOFTWARE: WORD PERFECT 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/459,101a

FILING DATE: June 2, 1995

CLASSIFICATION: 536

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/US94/07736

FILING DATE: 12 JUL 94

ATTORNEY/AGENT INFORMATION:

NAME: FERRARO, GREGORY D.

REGISTRATION NUMBER: 36,134

REFERENCE/DOCKET NUMBER: 325800-317

TELECOMMUNICATION INFORMATION:

TELEPHONE: 201-994-1700

TELEFAX: 201-994-1744

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 375 AMINO ACIDS

TYPE: AMINO ACID

STRANDEDNESS:

TOPOLOGY: LINEAR

MOLECULE TYPE: PROTEIN
US-08-459-101a-2

Query Match 92.8%; Score 1964.5; DB 2; Length 375;
Best Local Similarity 95.5%; Pred. No. 1.3e-163;
Matches 357; Conservative 2; Mismatches 14; Indels 1; Gaps 1;

```

QY 1 MSSRRALAVLTLHLTRALSTCPACACPLEAPRCAPGVGLVRGCGCCCKVCAROL 60
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 MSSRRALAVLTLHLTRVGLSTCPADCHCPLAPRCAPGVGLVRGCGCCCKVCAROL 60
QY 61 NEDSKTOPCDHTKGLKLECNFGASSSTALGICRAOSEGRPCENSRITYNGESFQPCNOHQ 120
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 NEDCKTOPCDHTKGLKLECNFGASSSTALGICRAOSEGRPCENSRITYNGESFQPCNOHQ 120
QY 121 CTCIDGANG-CIPCLPOELSLPNLGCNPNRLVKTGQCCCEWVCDEDSIKDPMEDQDGL 179
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 121 CTCIGWRGACIPLCPQELSLPNLGCNPNRLVKTGQCCCEWVCDEDSIKDPMEDQDGL 180
QY 180 GKELGFDSAEVELTRNNELIANGKRSILKRLPVFGMEPRILYNPLQGGKCIYQTSWSQC 239
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 181 GKELGFDSAEVELTRNNELIANGKRSILKRLPVFGMEPRILYNPLQGGKCIYQTSWSQC 240
QY 240 SKTCGTGISTVTNDNPECRVLKETRICEVRPCGQPVYSSILKKGKSKTKKSPPEVRF 299
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 241 SKTCGTGISTVTNDNPECRVLKETRICEVRPCGQPVYSSILKKGKSKTKKSPPEVRF 300
QY 300 YAGCLSVKRYPRKYGSCVDGRCTPOLTRTVKMFRCDEDETSKNVMYIOSCKCNYNC 359
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 301 YAGCLSVKRYPRKYGSCVDGRCTPOLTRTVKMFRCDEDETSKNVMYIOSCKCNYNC 360
QY 360 PHANEAAPFYRLF 373
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 361 PHANEAAPFYRLF 374

```

RESULT 5

US-08-468-847B-11
Sequence 11, Application US/08468847B

Patent No. 5780263

GENERAL INFORMATION:

APPLICANT: Hastings, Gregg A. and Adams, Mark D.

TITLE OF INVENTION: Human CCN-Like Growth Factor

NUMBER OF SEQUENCES: 20

CORRESPONDENCE ADDRESS:

ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,

ADDRESSEE: CECCHI, STEWART & OLSTEIN

STREET: 6 BECKER FARM ROAD

CITY: ROSELAND

STATE: NEW JERSEY

COUNTRY: USA

ZIP: 07068

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 INCH DISKETTE

COMPUTER: IBM PS/2

OPERATING SYSTEM: MS-DOS

SOFTWARE: WORD PERFECT 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/468,847B

FILING DATE: 6 June 1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: MULLINS, J.G.

REGISTRATION NUMBER: 33,073

REFERENCE/DOCKET NUMBER: 325800-442

TELECOMMUNICATION INFORMATION:

TELEPHONE: 201-994-1700

TELEFAX: 201-994-1744

INFORMATION FOR SEQ ID NO: 11:

SEQUENCE CHARACTERISTICS:

LENGTH: 379 AMINO ACIDS
 TYPE: AMINO ACID
 STRANDEDNESS:
 TOPOLOGY: LINEAR
 MOLECULE TYPE: PROTEIN
 US-08-468-847B-11

Query Match 91.2%; Score 1929; DB 1; Length 379;
 Best Local Similarity 90.9%; Pred. No. 1,6e-160;
 Matches 348; Conservative 10; Mismatches 19; Indels 6; Gaps 2;

QY 1 MSSRIARALAVLTLLHLRLALSTCPACCHCPLEAPKAPGVGLVYDGGCCCKVCAKOL 60
 DB 1 MSSSTRITLAVATLTLHLRLALSTCPACCHCPLEAPKAPGVGLVYDGGCCCKVCAKOL 60
 QY 61 NEDCSKTOPCDHRTKGLKGCNFGASSTALKGICRAOSBGRCEYNSRIYONGESFOPNCKHQ 120
 DB 61 NEDCSKTOPCDHRTKGLKGCNFGASSTALKGICRAOSBGRCEYNSRIYONGESFOPNCKHQ 120
 QY 121 CTCIDAVGCIPLCPDELSPNLGCPNPLVYKYGOCCEWVDEDSIKDPMEDDGLG 180
 DB 121 CTCIDAVGCIPLCPDELSPNLGCPNPLVYKYGOCCEWVDEDSIKDPMEDDGLG 180
 QY 181 KELGFASVEELTRNNELTAVGKRSKRLPVGMPRIYNYPL--QGOKCIYQTSMSQ 238
 DB 179 --LGIDASEVELTRNNELTAVGKSSSLKRLPVGTEPRVLFNPLHAHGOKCIYQTSMSQ 236
 QY 239 CSKTCGTGISTRYTNDNPECLVKEIRICEVRPCGQPVYSSLKKGKCKSTKKSPEPVRF 298
 DB 237 CSKSCGTGISTRYTNDNPECLVKEIRICEVRPCGQPVYSSLKKGKCKSTKKSPEPVRF 296
 QY 299 TYAGCISVKKYRKRYGSCVDGRCCTPLQTRIVKMFRCDEGEMFSKNVMIOSCKCNYN 358
 DB 297 TYAGCISVKKYRKRYGSCVDGRCCTPLQTRIVKMFRCDEGEMFSKNVMIOSCKCNYN 356
 QY 359 CPHNEAPFPYRLFNDIHKFRD 381
 DB 357 CPHNEASFRLYSLFNDIHKFRD 379

RESULT 6
 US-09-142-569-2
 ; Sequence 2, Application US/09142569
 ; Patent No. 6413735
 ; GENERAL INFORMATION:
 ; APPLICANT: Lau, Lester F.
 ; TITLE OF INVENTION: Extracellular Matrix Signalling Molecules
 ; NUMBER OF SEQUENCES: 17
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
 ; STREET: 6300 Sears Tower, 233 South Wacker Drive
 ; CITY: Chicago
 ; STATE: Illinois
 ; COUNTRY: United States of America
 ; ZIP: 60606-6402
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/142,569
 ; FILING DATE: 02-Apr-1999
 ; CLASSIFICATION: <Unknown>
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Clough, David W.
 ; REGISTRATION NUMBER: 36,107
 ; REFERENCE/DOCKET NUMBER: 28758/33766
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 312/474-6300
 ; TELEFAX: 312/474-0448
 ; TELEX: 25-3856
 ; INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:
 LENGTH: 379 amino acids
 TYPE: amino acid
 TOPOLOGY: Linear
 MOLECULE TYPE: protein
 FEATURE:
 NAME/KEY: misc-feature
 OTHER INFORMATION: "Mouse Cyt61 amino acid sequence"
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
 US-09-142-569-2

QY 1 MSSRIARALAVLTLLHLRLALSTCPACCHCPLEAPKAPGVGLVYDGGCCCKVCAKOL 60
 DB 1 MSSSTRITLAVATLTLHLRLALSTCPACCHCPLEAPKAPGVGLVYDGGCCCKVCAKOL 60
 QY 61 NEDCSKTOPCDHRTKGLKGCNFGASSTALKGICRAOSBGRCEYNSRIYONGESFOPNCKHQ 120
 DB 61 NEDCSKTOPCDHRTKGLKGCNFGASSTALKGICRAOSBGRCEYNSRIYONGESFOPNCKHQ 120
 QY 121 CTCIDAVGCIPLCPDELSPNLGCPNPLVYKYGOCCEWVDEDSIKDPMEDDGLG 180
 DB 121 CTCIDAVGCIPLCPDELSPNLGCPNPLVYKYGOCCEWVDEDSIKDPMEDDGLG 180
 QY 181 KELGFASVEELTRNNELTAVGKRSKRLPVGMPRIYNYPL--QGOKCIYQTSMSQ 238
 DB 179 --LGIDASEVELTRNNELTAVGSSSLKRLPVGTEPRVLFNPLHAHGOKCIYQTSMSQ 236
 QY 239 CSKTCGTGISTRYTNDNPECLVKEIRICEVRPCGQPVYSSLKKGKCKSTKKSPEPVRF 298
 DB 237 CSKSCGTGISTRYTNDNPECLVKEIRICEVRPCGQPVYSSLKKGKCKSTKKSPEPVRF 296
 QY 299 TYAGCISVKKYRKRYGSCVDGRCCTPLQTRIVKMFRCDEGEMFSKNVMIOSCKCNYN 358
 DB 297 TYAGCISVKKYRKRYGSCVDGRCCTPLQTRIVKMFRCDEGEMFSKNVMIOSCKCNYN 356
 QY 359 CPHNEAPFPYRLFNDIHKFRD 381
 DB 357 CPHNEASFRLYSLFNDIHKFRD 379

RESULT 7
 US-08-468-847B-13
 ; Sequence 13, Application US/08468847B
 ; Patent No. 5780263
 ; GENERAL INFORMATION:
 ; APPLICANT: Hastings, Gregg A. and Adams, Mark D.
 ; TITLE OF INVENTION: Human CCN-Like Growth Factor
 ; NUMBER OF SEQUENCES: 20
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLIAN,
 ; STREET: 6 BECKER FARM ROAD
 ; CITY: ROSELAND
 ; STATE: NEW JERSEY
 ; COUNTRY: USA
 ; ZIP: 07068
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: 3.5 INCH DISKETTE
 ; COMPUTER: IBM PS/2
 ; OPERATING SYSTEM: MS-DOS
 ; SOFTWARE: WORD PERFECT 5.1
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/468,847B
 ; FILING DATE: 6 June 1995
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:

NAME: MULLINS, J.G.
REGISTRATION NUMBER: 33, 073
REFERENCE/DOCKET NUMBER: 325800-442
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 375 AMINO ACIDS
TYPE: AMINO ACID
STRANDEDNESS:
TOPOLOGY: LINEAR
MOLECULE TYPE: PROTEIN
US-08-468-847B-13

Query Match 80.1%; Score 1695; DB 1; Length 375;
Best Local Similarity 81.0%; Pred. No. 4,2e-140;
Matches 311; Conservative 20; Mismatches 41; Indels 12; Gaps 7;

OY 1 MSRIARALAVTLHLTRLAL-STCPACHCPLEAKCAPGVGLVDRGGCCCKVCAKQ 59
DB 1 MGSAGARP-ALAAALICLARIALGSPCAVOCPPAAPQCAPGVGLVDRGGCCCKVCAKQ 59
OY 60 LNEDESKTOPCDHRTKGLCECNFGASSTALKGICRAQSEGRPCENSRITYONGESEFPQNCQ 119
DB 60 LNEDESKTOPCDHRTKGLCECNFGASSTALKGICRAQSEGRPCENSRITYONGESEFPQNCQ 119
OY 120 QCTCIDAVGCIPLCPQELSLPNLGCPNRLVYKVTGCCCEWVCDEDSIKDPMEDQGL 179
DB 120 QCTCIDAVGCIPLCPQELSLPNLGCPNRLVYKVTGCCCEWVCDEDSIKDPMEDQGL 179
OY 180 GKELGFASVEELRNNEELIIVAGKGRSLRLPVGMEPR--RLTVNPLDQOKCIYOTTSMS 237
DB 178 SKREGLDASEELRNNEELIIVAGKGRSLRLPVGMEPR--RLTVNPLDQOKCIYOTTSMS 237
OY 238 QCSKTGISTRTVNDNPECKLVKTRICEVRPCGQVYSSLSKKKKSKTKKSPPEVR 297
DB 238 QCSKTGISTRTVNDNPECKLVKTRICEVRPCGQVYSSLSKKKKSKTKKSPPEVR 297
OY 298 FTYAGCLSVKRYRKYRGCSYVDRGCTPQTLRTYKMRRCDEGEFTSKNVMIIOSCKNY 357
DB 298 FTYAGCLSVKRYRKYRGCSYVDRGCTPQTLRTYKMRRCDEGEFTSKNVMIIOSCKNY 357
OY 358 NCPHANEAFPEYRLFNDFIHKFRD 381
DB 358 NCPHANEAFPEYRLFNDFIHKFRD 381

RESULT 8
US-09-582-337-2
Sequence 2, Application US/09582337

PATENT NO. 6562618
GENERAL INFORMATION:
APPLICANT: Japan Tobacco, Inc.
TITLE OF INVENTION: Monoclonal Antibody Against Connective Tissue Growth Factor
TITLE OF INVENTION: and Medicinal Uses Thereof
FILE REFERENCE: JI-009PCT
CURRENT APPLICATION NUMBER: US/09/582,337
CURRENT FILING DATE: 2000-06-23
PRIOR APPLICATION NUMBER: JP P1997-367699
PRIOR FILING DATE: 1997-12-25
PRIOR APPLICATION NUMBER: JP P1998-356183
PRIOR FILING DATE: 1998-12-15
NUMBER OF SEQ ID NOS: 27
SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 2
LENGTH: 347
TYPE: PRT
ORGANISM: Rat
US-09-582-337-2

Query Match 45.0%; Score 953; DB 4; Length 347;
Best Local Similarity 46.2%; Pred. No. 2,2e-75;

Matches 176; Conservative 58; Mismatches 107; Indels 40; Gaps 7;

OY 1 MSRIARALAVTLHLTRLAL-STCPACHCPLEAKCAPGVGLVDRGGCCCKVCAKQ 58
DB 1 MGSAGARP-ALAAALICLARIALGSPCAVOCPPAAPQCAPGVGLVDRGGCCCKVCAKQ 60
OY 59 QUNEDSKTOPCDHRTKGLCECNFGASSTALKGICRAQSEGRPCENSRITYONGESEFPQNCQ 118
DB 61 QUNEDSKTOPCDHRTKGLCECNFGASSTALKGICRAQSEGRPCENSRITYONGESEFPQNCQ 119
OY 119 HOCTCIDAVGCIPLCPQELSLPNLGCPNRLVYKVTGCCCEWVCDEDSIKDPMEDQGL 178
DB 120 HOCTCIDAVGCIPLCPQELSLPNLGCPNRLVYKVTGCCCEWVCDEDSIKDPMEDQGL 178
OY 179 GKELGFASVEELRNNEELIIVAGKGRSLRLPVGMEPR--RLTVNPLDQOKCIYOTTSMS 237
DB 168 GKELGFASVEELRNNEELIIVAGKGRSLRLPVGMEPR--RLTVNPLDQOKCIYOTTSMS 205
OY 238 QCSKTGISTRTVNDNPECKLVKTRICEVRPCGQVYSSLSKKKKSKTKKSPPEVR 297
DB 206 ACSKTGISTRTVNDNPECKLVKTRICEVRPCGQVYSSLSKKKKSKTKKSPPEVR 265
OY 298 FTYAGCLSVKRYRKYRGCSYVDRGCTPQTLRTYKMRRCDEGEFTSKNVMIIOSCKNY 357
DB 266 FTYAGCLSVKRYRKYRGCSYVDRGCTPQTLRTYKMRRCDEGEFTSKNVMIIOSCKNY 325
OY 358 NCPHANEAFPEYRLFNDFIHKFRD 376
DB 326 NCPHANEAFPEYRLFNDFIHKFRD 346

RESULT 9
US-09-292-036-3
Sequence 3, Application US/09292036

PATENT NO. 6358741
GENERAL INFORMATION:
APPLICANT: FIBROGEN, INC.
APPLICANT: SCHMIDT, Brian
APPLICANT: ALLEN, Margaret
APPLICANT: SVERDRUP, Fran
TITLE OF INVENTION: CONNECTIVE TISSUE GROWTH FACTOR (CTGF) AND METHODS OF USE
FILE REFERENCE: FIBRO100-1
CURRENT APPLICATION NUMBER: US 09/292,036
CURRENT FILING DATE: 1999-04-14
PRIOR APPLICATION NUMBER: US 09/292,036
PRIOR FILING DATE: 1999-04-14
PRIOR APPLICATION NUMBER: US 09/187,478
PRIOR FILING DATE: 1998-11-06
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn version 3.0
SEQ ID NO 3
LENGTH: 348
TYPE: PRT
ORGANISM: Mouse
US-09-292-036-3

Query Match 45.0%; Score 952.5; DB 4; Length 348;
Best Local Similarity 46.6%; Pred. No. 2,5e-75;
Matches 178; Conservative 57; Mismatches 106; Indels 41; Gaps 8;

OY 1 MSRIARALAVTLHLTRLAL-STCPACHCPLEAKCAPGVGLVDRGGCCCKVCAKQ 57
DB 1 MGSAGARP-ALAAALICLARIALGSPCAVOCPPAAPQCAPGVGLVDRGGCCCKVCAKQ 60
OY 58 QUNEDSKTOPCDHRTKGLCECNFGASSTALKGICRAQSEGRPCENSRITYONGESEFPQNCQ 117
DB 61 QUNEDSKTOPCDHRTKGLCECNFGASSTALKGICRAQSEGRPCENSRITYONGESEFPQNCQ 119
OY 118 HOCTCIDAVGCIPLCPQELSLPNLGCPNRLVYKVTGCCCEWVCDEDSIKDPMEDQGL 177
DB 120 HOCTCIDAVGCIPLCPQELSLPNLGCPNRLVYKVTGCCCEWVCDEDSIKDPMEDQGL 168

QY 178 LKGLKELGPASAEVELTRNNELIAVGKGRSLKRL-PVFGMEPRLLYNPLOGOKCIYOTISM 236
Db 169 -----KDRTAVGPAALAYRLIEDTFEGDPTMM-----RANLVQPTTEW 205
QY 237 SOCSTKCTGISTRTVNDNPECLVETRICERPCGQPVYSSLKGGKSKTKRSPEPV 296
Db 206 SACSCTGCGISIRTVNDNTPFCHLEKOSRLCWRPCPADLEENIKKKKCIKRPKIAKRV 265
QY 297 REFTYAGCLSVKKYRPYCGSCVDGRCTPOLRTVYMRPCDEDEFTFSNNVMIOGCKCN 356
Db 266 KEFLSGCTSVKTYRAKFCVCTDGRCTPHRTTLVEFKCPDGEIMKKNNMFIKTCACH 325
QY 357 YNCPHANEAFPEY--RLFNDI 376
Db 326 YNCPGNDLIFESLYRKMYGDM 347

RESULT 10

US-08-468-847B-15
Sequence 15, Application US/08468847B
Patent No. 5780263
GENERAL INFORMATION:
APPLICANT: Hastings, Gregg A. and Adams, Mark D.
TITLE OF INVENTION: Human CCN-like Growth Factor
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: CARIELLA, BYRNE, BAIN, GILFILLIAN,
ADDRESSEE: CECCHI, STEWART & OLSTEIN
STREET: 6 BECKER FARM ROAD
CITY: ROSELAND
STATE: NEW JERSEY
COUNTRY: USA
ZIP: 07068
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 INCH DISKETTE
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: WORD PERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/468,847B
FILING DATE: 6 June 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: MULLINS, J.G.
REGISTRATION NUMBER: 33,073
REFERENCE/DOCKET NUMBER: 325800-442
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 348 AMINO ACIDS
TYPE: AMINO ACID
STRANDEDNESS:
TOPOLOGY: LINEAR
MOLECULE TYPE: PROTEIN
US-08-468-847B-15

Query Match 44.8%; Score 948.5; DB 1; Length 348;
Best Local Similarity 46.3%; Pred. No. 5.5e-75;
Matches 177; Conservative 58; Mismatches 106; Indels 41; Gaps 8;
QY 1 MSSRIARALAVVTLHL-TRLAL-STCPAACHPLE-APKCAPGVGVRODGGCGCKYCA 57
Db 1 MLASVAGPISLALVILLALCTRPATGDCSAQCQCAAEAPHPCPAGVSLVLDGCGCRVCA 60
QY 58 KQLNDCSKTOPCDHTKGLCNFGASSTALNGICRAQSGRPECENSRITYQNGESFQPN 117
Db 61 KQLGELCTERDPCDHPKGLFDFGSPANRKIGVCIAK-DGAPCVFGSGSVYRSGESFQSSC 119

QY 118 QHOCITIDGAVGICPLCPDELSTPNIGCPNRPVYKYGQCCSEWVCDSDSIKDPMEDQG 177
Db 120 KYQCTCIDGAVGICPLCSMDVRLPSPDCEFPERRVKLPBGCKCKWVCDSD-----168
QY 178 LKGLKELGPASAEVELTRNNELIAVGKGRSLKRL-PVFGMEPRLLYNPLOGOKCIYOTISM 236
Db 169 -----KDRTAVGPAALAYRLIEDTFEGDPTMM-----RANLVQPTTEW 205
QY 237 SOCSTKCTGISTRTVNDNPECLVETRICERPCGQPVYSSLKGGKSKTKRSPEPV 296
Db 206 SACSCTGCGISIRTVNDNTPFCHLEKOSRLCWRPCPADLEENIKKKKCIKRPKIAKRV 265
QY 297 REFTYAGCLSVKKYRPYCGSCVDGRCTPOLRTVYMRPCDEDEFTFSNNVMIOGCKCN 356
Db 266 KEFLSGCTSVKTYRAKFCVCTDGRCTPHRTTLVEFKCPDGEIMKKNNMFIKTCACH 325
QY 357 YNCPHANEAFPEY--RLFNDI 376
Db 326 YNCPGNDLIFESLYRKMYGDM 347

RESULT 11

US-09-142-569-6
Sequence 6, Application US/09142569
Patent No. 6413735
GENERAL INFORMATION:
APPLICANT: Lau, Lester F.
TITLE OF INVENTION: Extracellular Matrix Signalling Molecules
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/142,569
FILING DATE: 02-Apr-1999
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Clough, David W.
REGISTRATION NUMBER: 36,107
REFERENCE/DOCKET NUMBER: 28758/33766
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 348 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: misc.feature
OTHER INFORMATION: "Fisp12 amino acid sequence"
SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-142-569-6

Query Match 44.8%; Score 948.5; DB 4; Length 348;
Best Local Similarity 46.3%; Pred. No. 5.5e-75;
Matches 177; Conservative 58; Mismatches 106; Indels 41; Gaps 8;
QY 1 MSSRIARALAVVTLHL-TRLAL-STCPAACHPLE-APKCAPGVGVRODGGCGCKYCA 57
Db 1 MLASVAGPISLALVILLALCTRPATGDCSAQCQCAAEAPHPCPAGVSLVLDGCGCRVCA 60


```

CORRESPONDENCE ADDRESS:
ADDRESS: Spensley Horn Judas & Lubitz
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: US
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/167,628
FILING DATE:
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US/07/752,427
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Wetherell, Jr. Ph.D., John W.
REGISTRATION NUMBER: 31,678
REFERENCE/DOCKET NUMBER: PD-1294
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-455-5100
TELEFAX: 619-455-5110
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 349 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-167-628-2

Query Match
Best Local Similarity 46.6%; Score 947; DB 1; Length 349;
Matches 174; Conservative 58; Mismatches 101; Indels 40; Gaps 8;

44.8%; Score 947; DB 1; Length 349;
Best Local Similarity 46.6%; Pred. No. 7.4e-75;
Matches 174; Conservative 58; Mismatches 101; Indels 40; Gaps 8;

9 TALLVTLHLRLAL-STCPAACHCPL-APRCAPGVGLVRDGGCCCKVCAKOLNEDCSK 66
11 VAFVLLALCSRPAGVQNSGPCRCPDEPAPRCAGVSLVLDGCCRCVCAKOLGELCTE 70
67 TOPCHTGLBGNFGASSALKGICRAOEGSRPCENSRITYNGSGFOPNCHOCCTIDG 126
71 RDCPDHGLFCDPFSPPARKIGVCTAK-DGAPCIJFGTVRSGESFSSCKYQCTCIDG 129
127 AVGCIPLPQELSLNLGCPNRLVKTGOCCEEWCDSDSIKDPMEDQDGLGKELGFD 186
130 AVGCMPLCSMDVRLSPDCPPFRVRKLPKGCCEEWCDSDSIKDPMEDQDGLGKELGFD 172
187 ASEVELTRNNELIAVKGSRSLKRL-PVFGMEPRILYNPLQOGKCIYVQTTSMSCSKTGT 245
173 -----TVGPAALAAVRLDFTGPDPTMI-----RANCLVQTTSMSCSKTGT 215
246 GISTRTVNDNPECRVYKERICVEVRPCGOPYSSLSKKGKSKTKKSPVAFYTAGCLS 305
216 GISTRTVNDNPECRVYKERICVEVRPCGOPYSSLSKKGKSKTKKSPVAFYTAGCLS 275
306 VKYRPKYGSCVDRCCTPOLRTVKKRFRCEDEFTSKNVMYIOSCKNYNCPHANEA 365
276 MKTYRAKFCGYCTDRCCCTPHRTTLLPYEFKCPDGEVAKKMMFIKTCAHNCFGDNDI 335
366 AFPEY--RLFNDI 376
336 FESLYRRKMYGDM 348

```

```

APPLICANT: Bradham Jr., Douglas M.,
TITLE OF INVENTION: CONNECTIVE TISSUE GROWTH FACTOR
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESS: Spensley Horn Judas & Lubitz
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: US
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/386,680
FILING DATE: 10-FEB-1995
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US/08/167,628
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Wetherell, Jr. Ph.D., John W.
REGISTRATION NUMBER: 31,678
REFERENCE/DOCKET NUMBER: PD-1294
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-455-5100
TELEFAX: 619-455-5110
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 349 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-386-680-2

Query Match
Best Local Similarity 46.6%; Score 947; DB 1; Length 349;
Matches 174; Conservative 58; Mismatches 101; Indels 40; Gaps 8;

44.8%; Score 947; DB 1; Length 349;
Best Local Similarity 46.6%; Pred. No. 7.4e-75;
Matches 174; Conservative 58; Mismatches 101; Indels 40; Gaps 8;

9 TALLVTLHLRLAL-STCPAACHCPL-APRCAPGVGLVRDGGCCCKVCAKOLNEDCSK 66
11 VAFVLLALCSRPAGVQNSGPCRCPDEPAPRCAGVSLVLDGCCRCVCAKOLGELCTE 70
67 TOPCHTGLBGNFGASSALKGICRAOEGSRPCENSRITYNGSGFOPNCHOCCTIDG 126
71 RDCPDHGLFCDPFSPPARKIGVCTAK-DGAPCIJFGTVRSGESFSSCKYQCTCIDG 129
127 AVGCIPLPQELSLNLGCPNRLVKTGOCCEEWCDSDSIKDPMEDQDGLGKELGFD 186
130 AVGCMPLCSMDVRLSPDCPPFRVRKLPKGCCEEWCDSDSIKDPMEDQDGLGKELGFD 172
187 ASEVELTRNNELIAVKGSRSLKRL-PVFGMEPRILYNPLQOGKCIYVQTTSMSCSKTGT 245
173 -----TVGPAALAAVRLDFTGPDPTMI-----RANCLVQTTSMSCSKTGT 215
246 GISTRTVNDNPECRVYKERICVEVRPCGOPYSSLSKKGKSKTKKSPVAFYTAGCLS 305
216 GISTRTVNDNPECRVYKERICVEVRPCGOPYSSLSKKGKSKTKKSPVAFYTAGCLS 275
306 VKYRPKYGSCVDRCCTPOLRTVKKRFRCEDEFTSKNVMYIOSCKNYNCPHANEA 365
276 MKTYRAKFCGYCTDRCCCTPHRTTLLPYEFKCPDGEVAKKMMFIKTCAHNCFGDNDI 335
366 AFPEY--RLFNDI 376
336 FESLYRRKMYGDM 348

```

Wed Aug 6 07:53:12 2003

Job time : 15.0368 secs

us-09-495-448a-4.rai

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 5, 2003, 14:06:57 ; Search time 34.5908 Seconds

(without alignments)
1308.079 Million cell updates/sec

Title: US-09-495-448a-4

Perfect score: 2116
Sequence: 1 MSSRIARALALVLTILHLTR.....ANEAPFFYRLPNDIHKFRD 381

Scoring table:

BIOSUM62
Gapop 10.0 , Gapept 0.5

Searched: 451899 seqs, 118759770 residues

Total number of hits satisfying chosen parameters: 451899

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Published_Applications_AA:*

- 1: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/1/pubppaa/PCR_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/1/pubppaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/1/pubppaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/1/pubppaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/1/pubppaa/PCRUS_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/1/pubppaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/1/pubppaa/US09C_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/1/pubppaa/US09C_PUBCOMB.pep:*
- 13: /cgn2_6/ptodata/1/pubppaa/US10A_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/1/pubppaa/US10B_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/1/pubppaa/US10C_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/1/pubppaa/US10_NEW_PUB.pep:*
- 17: /cgn2_6/ptodata/1/pubppaa/US60_NEW_PUB.pep:*
- 18: /cgn2_6/ptodata/1/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	2116	100.0	381 14	US-10-053-753-4	Sequence 4, Appl1
2	2106	99.5	381 11	US-09-901-910-2	Sequence 2, Appl1
3	2106	99.5	381 15	US-10-294-796-2	Sequence 2, Appl1
4	2106	99.5	455 9	US-09-925-101-1432	Sequence 1432, Ap
5	2098	99.2	381 15	US-10-205-823-84	Sequence 84, Appl
6	1971.5	93.2	374 9	US-09-853-6258-12	Sequence 12, Appl
7	1971.5	93.2	375 11	US-09-901-910-7	Sequence 7, Appl1
8	1929	91.2	379 14	US-09-853-6258-11	Sequence 11, Appl
9	1929	91.2	375 9	US-10-053-753-2	Sequence 13, Appl1
10	1695	80.1	347 15	US-09-853-6258-13	Sequence 7, Appl1
11	955	45.1	348 14	US-10-245-977-7	Sequence 3, Appl1
12	952.5	45.0	348 9	US-10-101-040-3	Sequence 15, Appl
13	948.5	44.8	348 14	US-09-853-6258-15	Sequence 6, Appl1
14	948.5	44.8	348 15	US-10-053-753-6	Sequence 8, Appl1
15	948.5	44.8	348 15	US-10-245-977-8	Sequence 8, Appl1

16	948	44.8	347 14	US-10-080-173-2	Sequence 2, Appl1
17	948	44.8	347 14	US-10-101-040-2	Sequence 2, Appl1
18	947	44.8	349 14	US-10-101-040-4	Sequence 4, Appl1
19	947	44.8	349 14	US-10-011-859-26	Sequence 26, Appl
20	947	44.8	349 15	US-10-053-753-8	Sequence 8, Appl1
21	947	44.8	349 15	US-10-060-036-173	Sequence 173, Appl
22	947	44.8	349 15	US-10-171-111-46	Sequence 46, Appl
23	947	44.8	349 15	US-10-205-823-78	Sequence 78, Appl
24	947	44.8	349 15	US-10-245-977-2	Sequence 2, Appl1
25	938.5	44.4	348 9	US-09-853-6258-14	Sequence 14, Appl
26	933	44.1	349 15	US-10-245-977-5	Sequence 5, Appl1
27	906	42.8	349 9	US-09-853-6258-16	Sequence 16, Appl
28	851.5	40.2	351 9	US-09-853-6258-17	Sequence 17, Appl
29	827.5	39.1	357 9	US-09-853-6258-17	Sequence 17, Appl
30	827.5	39.1	357 14	US-10-011-859-25	Sequence 25, Appl
31	767	36.2	367 14	US-10-001-054-50	Sequence 50, Appl
32	767	36.2	367 15	US-10-112-267-4	Sequence 4, Appl1
33	767	36.2	367 15	US-10-112-267-8	Sequence 8, Appl1
34	766	36.2	367 15	US-10-112-267-7	Sequence 7, Appl1
35	766	36.2	367 15	US-10-112-267-22	Sequence 22, Appl
36	760.5	35.9	345 15	US-10-112-267-3	Sequence 3, Appl1
37	760.5	35.9	345 15	US-10-112-267-6	Sequence 6, Appl1
38	759.5	35.9	345 15	US-10-112-267-5	Sequence 5, Appl1
39	759.5	35.9	345 15	US-10-112-267-21	Sequence 21, Appl
40	758	35.8	367 15	US-10-112-267-12	Sequence 12, Appl
41	754.5	35.7	345 15	US-10-112-267-11	Sequence 11, Appl
42	613	29.0	339 15	US-10-112-267-36	Sequence 36, Appl
43	613	29.0	354 14	US-10-011-859-2	Sequence 2, Appl1
44	613	29.0	354 15	US-10-112-267-37	Sequence 37, Appl1
45	612	28.9	339 15	US-10-112-267-32	Sequence 32, Appl1

ALIGNMENTS

RESULT 1
US-10-053-753-4
Sequence 4, Application US/10053753
Publication NO. US20020150986A1
GENERAL INFORMATION:
APPLICANT: Lau, Lester F.
TITLE OF INVENTION: Extracellular Matrix Signalling Molecules
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borum
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/053, 753
FILING DATE: 22-Jan-2002
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Clough, David W.
REGISTRATION NUMBER: 36,107
REFERENCE/DOCKET NUMBER: 28758/33766
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 381 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

FEATURE:
NAME/KEY: misc_feature
OTHER INFORMATION: "Human Cyt61 amino acid sequence"
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-10-053-753-4

Query Match 100.0%; Score 2116; DB 14; Length 381;
Best Local Similarity 100.0%; Pred. No. 2.9e-173;
Matches 381; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MSSRIARALALVYTLHLTRLALSTCPACCHCLEAPKCAPGVGLVDDGGCCCKVCAKOL 60
DB 1 MSSRIARALALVYTLHLTRLALSTCPACCHCLEAPKCAPGVGLVDDGGCCCKVCAKOL 60
QY 61 NEDCSKTQPCDHTKGLKGCNFGASSTALKGICRAQSGRPECEYNSRIYONGESFQPCQHQ 120
DB 61 NEDCSKTQPCDHTKGLKGCNFGASSTALKGICRAQSGRPECEYNSRIYONGESFQPCQHQ 120
QY 121 CTCIDGAVGCIPLCPQELSLPNIAGCPNRLVKTGGCCCEWVDEDSIKDPMEDDGLLG 180
DB 121 CTCIDGAVGCIPLCPQELSLPNIAGCPNRLVKTGGCCCEWVDEDSIKDPMEDDGLLG 180
QY 181 KELGPDASEVELTRNNELLAVGKGRSLKRLPVFGMEPRILYNPLOGOKCIYVOTTSWQCS 240
DB 181 KELGPDASEVELTRNNELLAVGKGRSLKRLPVFGMEPRILYNPLOGOKCIYVOTTSWQCS 240
QY 241 KTCGTGISTRTVNDNECRVLKETRICEVPRPGOPVYSSILKKGKCKSKTKKSPPEVRFY 300
DB 241 KTCGTGISTRTVNDNECRVLKETRICEVPRPGOPVYSSILKKGKCKSKTKKSPPEVRFY 300
QY 301 AGCLSVKRYKPKYKGCSCVDGRCTPOLTRTVKMRFCEDGETFSKNVMMIQSCKNYNC 360
DB 301 AGCLSVKRYKPKYKGCSCVDGRCTPOLTRTVKMRFCEDGETFSKNVMMIQSCKNYNC 360
QY 361 HANEAAFPYRLFNDIHKFRD 381
DB 361 HANEAAFPYRLFNDIHKFRD 381
```

RESULT 2
US-09-901-910-2
; Sequence 2, Application US/09901910
; Publication No. US20030012768A1
; GENERAL INFORMATION:
; APPLICANT: Li, Haodong
; APPLICANT: Adams, Mark
; APPLICANT: Calenda Valerie
; TITLE OF INVENTION: Connective Tissue Growth Factor-2
; FILE REFERENCE: PF126P2
; CURRENT APPLICATION NUMBER: US/09/901,910
; CURRENT FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: 09/348,815
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: 08/459,101
; PRIOR FILING DATE: 1995-06-02
; PRIOR APPLICATION NUMBER: PCT/US94/07736
; PRIOR FILING DATE: 1994-07-12
; PRIOR APPLICATION NUMBER: 60/217,402
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/291,642
; PRIOR FILING DATE: 2001-05-18
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.0
; LENGTH: 381
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-901-910-2

Query Match 99.5%; Score 2106; DB 11; Length 381;
Best Local Similarity 99.5%; Pred. No. 2.1e-172;
Matches 379; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 MSSRIARALALVYTLHLTRLALSTCPACCHCLEAPKCAPGVGLVDDGGCCCKVCAKOL 60
DB 1 MSSRIARALALVYTLHLTRLALSTCPACCHCLEAPKCAPGVGLVDDGGCCCKVCAKOL 60
QY 61 NEDCSKTQPCDHTKGLKGCNFGASSTALKGICRAQSGRPECEYNSRIYONGESFQPCQHQ 120
DB 61 NEDCSKTQPCDHTKGLKGCNFGASSTALKGICRAQSGRPECEYNSRIYONGESFQPCQHQ 120
QY 121 CTCIDGAVGCIPLCPQELSLPNIAGCPNRLVKTGGCCCEWVDEDSIKDPMEDDGLLG 180
DB 121 CTCIDGAVGCIPLCPQELSLPNIAGCPNRLVKTGGCCCEWVDEDSIKDPMEDDGLLG 180
QY 181 KELGPDASEVELTRNNELLAVGKGRSLKRLPVFGMEPRILYNPLOGOKCIYVOTTSWQCS 240
DB 181 KELGPDASEVELTRNNELLAVGKGRSLKRLPVFGMEPRILYNPLOGOKCIYVOTTSWQCS 240
QY 241 KTCGTGISTRTVNDNECRVLKETRICEVPRPGOPVYSSILKKGKCKSKTKKSPPEVRFY 300
DB 241 KTCGTGISTRTVNDNECRVLKETRICEVPRPGOPVYSSILKKGKCKSKTKKSPPEVRFY 300
QY 301 AGCLSVKRYKPKYKGCSCVDGRCTPOLTRTVKMRFCEDGETFSKNVMMIQSCKNYNC 360
DB 301 AGCLSVKRYKPKYKGCSCVDGRCTPOLTRTVKMRFCEDGETFSKNVMMIQSCKNYNC 360
QY 361 HANEAAFPYRLFNDIHKFRD 381
DB 361 HANEAAFPYRLFNDIHKFRD 381
```

RESULT 3
US-10-294-796-2
; Sequence 2, Application US/10294796
; Publication No. US20030078391A1
; GENERAL INFORMATION:
; APPLICANT: Li, Haodong et al.
; TITLE OF INVENTION: Connective Tissue Growth Factor-2
; FILE REFERENCE: PF126PID2
; CURRENT APPLICATION NUMBER: US/10/294,796
; CURRENT FILING DATE: 2002-11-15
; PRIOR APPLICATION NUMBER: US 09/348,815
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: US 08/459,101
; PRIOR FILING DATE: 1995-06-02
; PRIOR APPLICATION NUMBER: PCT/US94/07736
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; LENGTH: 381
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-294-796-2

Query Match 99.5%; Score 2106; DB 15; Length 381;
Best Local Similarity 99.5%; Pred. No. 2.1e-172;
Matches 379; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 MSSRIARALALVYTLHLTRLALSTCPACCHCLEAPKCAPGVGLVDDGGCCCKVCAKOL 60
DB 1 MSSRIARALALVYTLHLTRLALSTCPACCHCLEAPKCAPGVGLVDDGGCCCKVCAKOL 60
QY 61 NEDCSKTQPCDHTKGLKGCNFGASSTALKGICRAQSGRPECEYNSRIYONGESFQPCQHQ 120
DB 61 NEDCSKTQPCDHTKGLKGCNFGASSTALKGICRAQSGRPECEYNSRIYONGESFQPCQHQ 120
QY 121 CTCIDGAVGCIPLCPQELSLPNIAGCPNRLVKTGGCCCEWVDEDSIKDPMEDDGLLG 180
DB 121 CTCIDGAVGCIPLCPQELSLPNIAGCPNRLVKTGGCCCEWVDEDSIKDPMEDDGLLG 180
QY 181 KELGPDASEVELTRNNELLAVGKGRSLKRLPVFGMEPRILYNPLOGOKCIYVOTTSWQCS 240
DB 181 KELGPDASEVELTRNNELLAVGKGRSLKRLPVFGMEPRILYNPLOGOKCIYVOTTSWQCS 240
```

QY 241 KTCGTGISTRTVNDNPECRVETRICCEVRCGOPYVSSLKGGKCKSKTKKSPVAFETY 300
DB 241 KTCGTGISTRTVNDNPECRVETRICCEVRCGOPYVSSLKGGKCKSKTKKSPVAFETY 300
QY 301 AGCLSVKRYKRYKCGSCVDGRCTPOLTRTVKMRFRCEDEFTSKNMMIOSCKCYNCP 360
DB 301 AGCLSVKRYKRYKCGSCVDGRCTPOLTRTVKMRFRCEDEFTSKNMMIOSCKCYNCP 360
QY 361 HANEAFPEYRLFNDDIHKFRD 381
DB 361 HANEAFPEYRLFNDDIHKFRD 381

RESULT 4
US-09-925-301-1432
; Sequence 1432, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: P1106
; CURRENT APPLICATION NUMBER: US/09/925,301
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: Patent Ver. 2.0
; SEQ ID NO 1432
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-301-1432

Query Match 99.5%; Score 2106; DB 9; Length 455;
Best Local Similarity 99.5%; Pred. No. 2.5e-172;
Matches 379; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSSRIARALALVYTLHLTRIALSTCPACHCPLEAPKCAPGVGLVBDGGCCCKVCAKOL 60
DB 75 MSSRIARALALVYTLHLTRIALSTCPACHCPLEAPKCAPGVGLVBDGGCCCKVCAKOL 134
QY 61 NEDCSKTOPCDHRTGELCNFGASTALKGICRAOSEBPCEYNSRITQONESPOPNCKHQ 120
DB 135 NEDCSKTOPCDHRTGELCNFGASTALKGICRAOSEBPCEYNSRITQONESPOPNCKHQ 194
QY 121 CTCIDGAVGCIPLCPQELSLPNIAGCPNRLVKTGQCCCEMWCDEDSIKPMEDQDGLG 180
DB 195 CTCIDGAVGCIPLCPQELSLPNIAGCPNRLVKTGQCCCEMWCDEDSIKPMEDQDGLG 254
QY 181 KELGFDASEVELTRNNELIANGKRSILKRLPVFGEERILYNPLOGKCIYQTTSMGQS 240
DB 255 KELGFDASEVELTRNNELIANGKRSILKRLPVFGEERILYNPLOGKCIYQTTSMGQS 314
QY 241 KTCGTGISTRTVNDNPECRVETRICCEVRCGOPYVSSLKGGKCKSKTKKSPVAFETY 300
DB 315 KTCGTGISTRTVNDNPECRVETRICCEVRCGOPYVSSLKGGKCKSKTKKSPVAFETY 374
QY 301 AGCLSVKRYKRYKCGSCVDGRCTPOLTRTVKMRFRCEDEFTSKNMMIOSCKCYNCP 360
DB 375 AGCLSVKRYKRYKCGSCVDGRCTPOLTRTVKMRFRCEDEFTSKNMMIOSCKCYNCP 434
QY 361 HANEAFPEYRLFNDDIHKFRD 381
DB 435 HANEAFPEYRLFNDDIHKFRD 455

RESULT 5
US-10-205-823-84
; Sequence 84, Application US/10205823
; Patent No. US20030108963A1
; GENERAL INFORMATION:

APPLICANT: Schlegel, Robert
APPLICANT: Monahan, John E.
APPLICANT: Endege, Wilson O.
APPLICANT: Gannavara, Manjula
APPLICANT: Gorbacheva, Bella
APPLICANT: Hoersch, Sebastian
APPLICANT: Kamatkar, Shubhangi
APPLICANT: Monsey, Angela M.
APPLICANT: Glatz, Karen
APPLICANT: Zhao, Xumei
APPLICANT: Anderson, Dustin

TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
THERAPY OF PROSTATE CANCER
FILE REFERENCE: MRI-044
CURRENT APPLICATION NUMBER: US/10/205,823
PRIOR FILING DATE: 2002-07-25
PRIOR APPLICATION NUMBER: 60/307,982
PRIOR FILING DATE: 2001-07-25
PRIOR APPLICATION NUMBER: 60/314,356
PRIOR FILING DATE: 2001-08-22
PRIOR APPLICATION NUMBER: 60/325,020
PRIOR FILING DATE: 2001-09-25
PRIOR APPLICATION NUMBER: 60/341,746
PRIOR FILING DATE: 2001-12-12
PRIOR APPLICATION NUMBER: 60/362,158
PRIOR FILING DATE: 2002-03-05
NUMBER OF SEQ ID NOS: 455
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 84
LENGTH: 381
TYPE: PRT
ORGANISM: Homo sapiens
US-10-205-823-84

Query Match 99.1%; Score 2098; DB 15; Length 381;
Best Local Similarity 99.0%; Pred. No. 1e-171;
Matches 377; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MSSRIARALALVYTLHLTRIALSTCPACHCPLEAPKCAPGVGLVBDGGCCCKVCAKOL 60
DB 1 MSSRIARALALVYTLHLTRIALSTCPACHCPLEAPKCAPGVGLVBDGGCCCKVCAKOL 60
QY 61 NEDCSKTOPCDHRTGELCNFGASTALKGICRAOSEBPCEYNSRITQONESPOPNCKHQ 120
DB 61 NEDCSKTOPCDHRTGELCNFGASTALKGICRAOSEBPCEYNSRITQONESPOPNCKHQ 120
QY 121 CTCIDGAVGCIPLCPQELSLPNIAGCPNRLVKTGQCCCEMWCDEDSIKPMEDQDGLG 180
DB 121 CTCIDGAVGCIPLCPQELSLPNIAGCPNRLVKTGQCCCEMWCDEDSIKPMEDQDGLG 180
QY 181 KELGFDASEVELTRNNELIANGKRSILKRLPVFGEERILYNPLOGKCIYQTTSMGQS 240
DB 181 KELGFDASEVELTRNNELIANGKRSILKRLPVFGEERILYNPLOGKCIYQTTSMGQS 240
QY 241 KTCGTGISTRTVNDNPECRVETRICCEVRCGOPYVSSLKGGKCKSKTKKSPVAFETY 300
DB 241 KTCGTGISTRTVNDNPECRVETRICCEVRCGOPYVSSLKGGKCKSKTKKSPVAFETY 300
QY 301 AGCLSVKRYKRYKCGSCVDGRCTPOLTRTVKMRFRCEDEFTSKNMMIOSCKCYNCP 360
DB 301 AGCLSVKRYKRYKCGSCVDGRCTPOLTRTVKMRFRCEDEFTSKNMMIOSCKCYNCP 360
QY 361 HANEAFPEYRLFNDDIHKFRD 381
DB 361 HANEAFPEYRLFNDDIHKFRD 381

RESULT 6
US-09-853-625B-12
; Sequence 12, Application US/09853625B
; Patent No. US20020049304A1
; GENERAL INFORMATION:

APPLICANT: Hastings, Gregg A. and Adams, Mark D.
TITLE OF INVENTION: Human CCN-Like Growth Factor
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLIAN,
CECCHI, STEWART & OLSTEIN
STREET: 6 BECKER FARM ROAD
CITY: ROSELAND
STATE: NEW JERSEY
COUNTRY: USA
ZIP: 07068

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 INCH DISKETTE
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: WORD PERFECT 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/853,625B
FILING DATE: 14-May-2001
CLASSIFICATION: <Unknown>
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 09/053,587
FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: MULLINS, J.G.
REGISTRATION NUMBER: 33,073
REFERENCE/DOCKET NUMBER: 325800-442
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744

INFORMATION FOR SEQ ID NO: 12:

SEQUENCE CHARACTERISTICS:
LENGTH: 374 AMINO ACIDS
TYPE: AMINO ACID
STRANDEDNESS: <Unknown>
MOLECULE TYPE: PROTEIN
SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-09-853-625B-12

Query Match 93.2%; Score 1971.5; DB 9; Length 374;
Best Local Similarity 95.7%; Pred. No. 6,6e-161;
Matches 358; Conservative 2; Mismatches 13; Indels 1; Gaps 1;

QY 1 MSSRIARALAVYTLHLRLALSTCPAACHCPLEAPKCAPGVGLVRDGGCCCKVCAKOL 60
DB 1 MSSRIARELAVYTLHLRLVGLSTCPADCHCPLEAPKCAPGVGLVRDGGCCCKVCAKOL 60
QY 61 NECCSTQPCDHTKGLGECNFGASSTALKGICRAOSGRCPEYNSRIYONGESFQPCNKHQ 120
DB 61 NECCRTQPCDHTKGLGECNFGASSTALKGICRAOSGRCPEYNSRIYONGESFQPCNKHQ 120
QY 121 CTCIDGAVG-CIPLCQELSLPMLGCPNRLVKGCCCEWVCDEDSIKDPMEDDGLL 179
DB 121 CTCIGRRACIPLCQELSLPMLGCPNRLVKGCCCEWVCDEDSIKDPMEDDGLL 180
QY 180 GKEIGDASEVELTRNNELIANGKGRSLKRLPFYGMERILYNPLOGOKCIYQTTSMSC 239
DB 181 GKEIGDASEVELTRNNELIANGKGRSLKRLPFYGMERILYNPLOGOKCIYQTTSMSC 240
QY 240 SKTCGTGISTRYTNDNPECLRYKETRICVRCGQGVYSSLKGGKCSKTKKSPPEVRF 299
DB 241 SKTCGTGISTRYTNDNPECLRYKETRICVRCGQGVYSSLKGGKCSKTKKSPPEVRF 300
QY 300 YAGCLSVKRYRPYKGCYDGRCTPOLRTVYKMRFRCDGFTFSKNVMIOSSCKNYNC 359
DB 301 YAGCLSVKRYRPYKGCYDGRCTPOLRTVYKMRFRCDGFTFSKNVMIOSSCKNYNC 360
QY 360 PHANEAAPPEYRLF 373
DB 361 PHANEAAPPEYRLF 374

RESULT 7
US-09-901-910-7
Sequence 7, Application US/09901910
Publication No. US20030012768A1
GENERAL INFORMATION:
APPLICANT: LI, Haodong
APPLICANT: Adams, Mark
TITLE OF INVENTION: Connective Tissue Growth Factor-2
FILE REFERENCE: PFI26P2
CURRENT APPLICATION NUMBER: US/09/901,910
CURRENT FILING DATE: 2001-07-11
PRIOR APPLICATION NUMBER: 09/348,815
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: 08/459,101
PRIOR FILING DATE: 1995-06-02
PRIOR APPLICATION NUMBER: PCT/US94/07736
PRIOR FILING DATE: 1994-07-12
PRIOR APPLICATION NUMBER: 60/217,402
PRIOR FILING DATE: 2000-07-11
PRIOR APPLICATION NUMBER: 60/291,642
PRIOR FILING DATE: 2001-05-18
NUMBER OF SEQ ID NOS: 8
SOFTWARE: PatentIn version 3.0
SEQ ID NO: 7
LENGTH: 375
TYPE: PRT
ORGANISM: homo sapiens
US-09-901-910-7

Query Match 93.2%; Score 1971.5; DB 11; Length 375;
Best Local Similarity 95.7%; Pred. No. 6,7e-161;
Matches 358; Conservative 2; Mismatches 13; Indels 1; Gaps 1;

QY 1 MSSRIARALAVYTLHLRLALSTCPAACHCPLEAPKCAPGVGLVRDGGCCCKVCAKOL 60
DB 1 MSSRIARELAVYTLHLRLVGLSTCPADCHCPLEAPKCAPGVGLVRDGGCCCKVCAKOL 60
QY 61 NECCSTQPCDHTKGLGECNFGASSTALKGICRAOSGRCPEYNSRIYONGESFQPCNKHQ 120
DB 61 NECCRTQPCDHTKGLGECNFGASSTALKGICRAOSGRCPEYNSRIYONGESFQPCNKHQ 120
QY 121 CTCIDGAVG-CIPLCQELSLPMLGCPNRLVKGCCCEWVCDEDSIKDPMEDDGLL 179
DB 121 CTCIGRRACIPLCQELSLPMLGCPNRLVKGCCCEWVCDEDSIKDPMEDDGLL 180
QY 180 GKEIGDASEVELTRNNELIANGKGRSLKRLPFYGMERILYNPLOGOKCIYQTTSMSC 239
DB 181 GKEIGDASEVELTRNNELIANGKGRSLKRLPFYGMERILYNPLOGOKCIYQTTSMSC 240
QY 240 SKTCGTGISTRYTNDNPECLRYKETRICVRCGQGVYSSLKGGKCSKTKKSPPEVRF 299
DB 241 SKTCGTGISTRYTNDNPECLRYKETRICVRCGQGVYSSLKGGKCSKTKKSPPEVRF 300
QY 300 YAGCLSVKRYRPYKGCYDGRCTPOLRTVYKMRFRCDGFTFSKNVMIOSSCKNYNC 359
DB 301 YAGCLSVKRYRPYKGCYDGRCTPOLRTVYKMRFRCDGFTFSKNVMIOSSCKNYNC 360
QY 360 PHANEAAPPEYRLF 373
DB 361 PHANEAAPPEYRLF 374

RESULT 8
US-09-853-625B-11
Sequence 11, Application US/09853625B
Patent No. US20020049304A1
GENERAL INFORMATION:
APPLICANT: Hastings, Gregg A. and Adams, Mark D.
TITLE OF INVENTION: Human CCN-Like Growth Factor
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLIAN,


```

US-09-853-625B-13
; Sequence 13, Application us/09853625B
; Patent No. US20020049304A1
;
GENERAL INFORMATION:
; APPLICANT: Hastings, Gregg A. and Adams, Mark D.
; TITLE OF INVENTION: Human CCN-Like Growth Factor
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLIAN,
; CROCHT, STEWART & OLSTEIN
; STREET: 6 BECKER FARM ROAD
; CITY: ROSELAND
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07068
;
COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: us/09/853,625B
; FILING DATE: 14-May-2001
; CLASSIFICATION: <Unknown>
;
PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/053,587
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: MULLINS, J.G.
; REGISTRATION NUMBER: 33,073
; REFERENCE/DOCKET NUMBER: 325800-442
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
;
INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 375 AMINO ACIDS
; TYPE: AMINO ACID
; STRANDEDNESS: <Unknown>
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PROTEIN
; SEQUENCE DESCRIPTION: SEQ ID NO: 13:
US-09-853-625B-13

Query Match
Best Local Similarity 80.1%; Score 1695; DB 9; Length 375;
Matches 311; Conservative 20; Mismatches 41; Indels 12; Gaps 7;

QY 1 MSSRIARALAVVTLHLTRLAL-STCPAACHCPLAPKAPGVGLVRDGGCGCKVCARQ 59
DB 1 MSAGAGAP-ALAAALICLARLALGSPCAVCCPAAAPGCAAGVGLVDPDGGCGCKVCARQ 59
QY 60 LNEDESKTQPCDHTKLECNFGASSTALKGICRAOSEGRCEYNSRIYONGSEFQNCQ 119
DB 60 LNEDESKTQPCDHTKLECNFGASSTALKGICRAOSEGRCEYNSRIYONGSEFQNCQ 119
QY 120 QCTCIGAVGCIPLCPOELSLPMLGCPNPLVAVTVOCCCEWVCDDESDIKDPMEQDGL 179
DB 120 QCTCIGAVGCIPLCPOELSLPMLGCPNPLVAVTVOCCCEWVCDDESDIKDPMEQDGL 179
QY 180 GKELGPDASEVELTRNNELIAYGKGRSLKRLPVFGMEP--RLIYNPLDGGKCIYQTTSS 237
DB 178 SKFGLDASGELTRNNELIAYKG--GLKMLPVFGSEPSRAENP-----KCIYQTTSS 232
QY 238 QCSKTGTGTSTVNTDNEPCRLVETRICIENRPGQPYYSILKKGKSKTKKSPSPYR 297
DB 233 QCSKTGTGTSTVNTDNEPCRLVETRICIENRPGQPYYSILKKGKSKTKKSPSPYR 292
QY 298 FTYAGGLSVKYPKPGSGCVDRCTPOLTRTVKMRFRCEDEGTFSSKNVMTIOSCKNY 357
DB 293 FTYAGGSVAKYKPKPGSGCVDRCTPOLTRTVKIRFRODDEITFTKSVMTIOSCRNT 352
QY 358 NCPHANEAAFPFYRLFNDIHKFRD 381

```

```

DB 353 NCPHANE-AAFPFYRLVNDIHKFRD 375
||||| :|||||
RESULT 11
US-10-245-977-7
; Sequence 7, Application us/10245977
; Publication No. US20030113816A1
;
GENERAL INFORMATION:
; APPLICANT: Weitz, Stephen L
; APPLICANT: Usinger, William R
; TITLE OF INVENTION: METHODS OF ASSAYING CONNECTIVE TISSUE GROWTH FACTOR
; FILE REFERENCE: EP0812 US
; CURRENT APPLICATION NUMBER: us/10/245,977
; CURRENT FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: US 60/323,305
; PRIOR FILING DATE: 2001-09-18
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 347
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-245-977-7

Query Match
Best Local Similarity 45.1%; Score 955; DB 15; Length 347;
Matches 176; Conservative 58; Mismatches 107; Indels 40; Gaps 7;

QY 1 MSSRIARALAVVTLHLTRLAL-STCPAACHCPLAPKAPGVGLVRDGGCGCKVCARQ 58
DB 1 MASVAGPYSLAVLLCTRPATGDDCSAOCOCABABARCPAGVSLVLDGGCGCKVCARQ 60
QY 59 QLNEDSKTQPCDHTKLECNFGASSTALKGICRAOSEGRCEYNSRIYONGSEFQNCQ 118
DB 61 QLNEDSKTQPCDHTKLECNFGASSTALKGICRAOSEGRCEYNSRIYONGSEFQNCQ 119
QY 119 HQCTCIDGAVGCIPLCPOELSLPMLGCPNPLVAVTVOCCCEWVCDDESDIKDPMEQDGL 178
DB 120 YQCTCIDGAVGCIPLCPOELSLPMLGCPNPLVAVTVOCCCEWVCDDESDIKDPMEQDGL 167
QY 179 LGKELGPDASEVELTRNNELIAYGKGRSLKRL-PVFGMEPRLIYNPLDGGKCIYQTTSS 237
DB 168 -----KRTVVGPALAAVRLDFTGPDPTM-----RANCIYQTTSS 205
QY 238 QCSKTGTGTSTVNTDNEPCRLVETRICIENRPGQPYYSILKKGKSKTKKSPSPYR 297
DB 206 ACSKTGSGISTVNTDNEPCRLVETRICIENRPGQPYYSILKKGKSKTKKSPSPYR 265
QY 298 FTYAGGLSVKYPKPGSGCVDRCTPOLTRTVKMRFRCEDEGTFSSKNVMTIOSCKNY 357
DB 266 FELSGTSVYKYPKPGSGCVDRCTPOLTRTVKMRFRCEDEGTFSSKNVMTIOSCKNY 325
QY 358 NCPHANEAAFPFY--RLFNDI 376
DB 326 NCPGDNDIFESLYRRMYGDM 346

RESULT 12
US-10-101-040-3
; Sequence 3, Application us/10101040
; Publication No. US20020142353A1
;
GENERAL INFORMATION:
; APPLICANT: FIBROGEN, INC
; APPLICANT: SCHMIDT, Brian
; APPLICANT: ALLEN, Margaret
; APPLICANT: SVERDRUP, Fran
; APPLICANT: CARMICHUEL, David
; TITLE OF INVENTION: CONNECTIVE TISSUE GROWTH FACTOR (CTGF) AND METHODS OF
; FILE REFERENCE: FIBRO1100-1
; CURRENT APPLICATION NUMBER: US/10/101,040
; CURRENT FILING DATE: 2002-03-18

```

PRIOR APPLICATION NUMBER: 09/292,036
PRIOR FILING DATE: 1999-04-14
PRIOR APPLICATION NUMBER: US 09/292,036
PRIOR FILING DATE: 1999-04-14
PRIOR APPLICATION NUMBER: US 09/187,478
PRIOR FILING DATE: 1998-11-06
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn version 3.0
SEQ ID NO 3
LENGTH: 348
TYPE: PRT
ORGANISM: Mouse
US-10-101-040-3

Query Match 45.0%; Score 952.5; DB 14; Length 348;
Best Local Similarity 46.6%; Pred. No. 1.1e-73;
Matches 178; Conservative 57; Mismatches 106; Indels 41; Gaps 8;

QY 1 MSRIARALAVTLVTLHL-TRIAL-STCPACHCPLF-APKCAPGVGLVRDGGCCRYCA 57
DB 1 MIAVAGPISIALVIALCTRPATGDCSACQCAAEAPHCPRAGVSLVLDGCGCCRYCA 60
QY 58 KQLNEDSKTOPCDHTKLECNFASSTALKGICRAQSEGRPCYNSRIYONGESFOPNC 117
DB 61 KQELCTERDPCDPHKGLECFDGFSPANKRIGVCTAK-DGAPCVGSGSVYSGESFQSSC 119
QY 118 OHQCTCIGAVGCIPLCQELSLPRLGCPNRLVYVQCCCEWVCEDESDIKDPEDDGG 177
DB 120 KYQCTCLDGAAGVCPCLSMVRLPSDPPRRVRLPKQCKCEWVCEDESDIKDPEDDGG 168
QY 178 LIGKELGPDASEVELTRNNELIANGKGRSLKRL-PVFGMEPRILYNPLQOGKIYQTSW 236
DB 169 -----KORTAVGPALAAARLEDTFDPDTM-----RANCLVQTTW 205
QY 237 SQSKTCGTGISTRTVNDNPECRILVETRIEVRPCGOPYSSLKKGKSKTKSPBPV 296
DB 206 SACSCTGCMGISTRTVNDNPECRILVETRIEVRPCGOPYSSLKKGKSKTKSPBPV 265
QY 297 RFTYAGCLSVKRYKPGYCGSVDCRCCTPOLTRIVYKMFREDGETFSKNMMQOSCKN 356
DB 266 KFEISGCTSVKRYKPGYCGSVDCRCCTPOLTRIVYKMFREDGETFSKNMMQOSCKN 325
QY 357 YNCPHANEAFPEY--RLFNDI 376
DB 326 YNCPGNDIFESILYRKMYGDM 347

RESULT 13
US-09-853-625B-15
Sequence 15, Application US/09853625B
Patent No. US20020049304A1
GENERAL INFORMATION:
APPLICANT: Hastings, Gregg A. and Adams, Mark D.
TITLE OF INVENTION: Human CCN-Like Growth Factor
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESSES:
ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLIAN,
CECCHI, STEWART & OLSTEIN
STREET: 6 BECKER FARM ROAD
CITY: ROSELAND
STATE: NEW JERSEY
COUNTRY: USA
ZIP: 07068
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 INCH DISKETTE
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: WORD PERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/853,625B
FILING DATE: 14-May-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/053,587
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: MOLLINS, J.G.
REGISTRATION NUMBER: 33,073
REFERENCE/DOCKET NUMBER: 325800-442
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 348 AMINO ACIDS
TYPE: AMINO ACID
STRANDEDNESS: <Unknown>
TOPOLOGY: LINEAR
MOLECULE TYPE: PROTEIN
SEQUENCE DESCRIPTION: SEQ ID NO: 15:
US-09-853-625B-15

Query Match 44.8%; Score 948.5; DB 9; Length 348;
Best Local Similarity 46.3%; Pred. No. 2.5e-73;
Matches 177; Conservative 58; Mismatches 106; Indels 41; Gaps 8;

QY 1 MSRIARALAVTLVTLHL-TRIAL-STCPACHCPLF-APKCAPGVGLVRDGGCCRYCA 57
DB 1 MIAVAGPISIALVIALCTRPATGDCSACQCAAEAPHCPRAGVSLVLDGCGCCRYCA 60
QY 58 KQLNEDSKTOPCDHTKLECNFASSTALKGICRAQSEGRPCYNSRIYONGESFOPNC 117
DB 61 KQELCTERDPCDPHKGLECFDGFSPANKRIGVCTAK-DGAPCVGSGSVYSGESFQSSC 119
QY 118 OHQCTCIGAVGCIPLCQELSLPRLGCPNRLVYVQCCCEWVCEDESDIKDPEDDGG 177
DB 120 KYQCTCLDGAAGVCPCLSMVRLPSDPPRRVRLPKQCKCEWVCEDESDIKDPEDDGG 168
QY 178 LIGKELGPDASEVELTRNNELIANGKGRSLKRL-PVFGMEPRILYNPLQOGKIYQTSW 236
DB 169 -----KORTAVGPALAAARLEDTFDPDTM-----RANCLVQTTW 205
QY 237 SQSKTCGTGISTRTVNDNPECRILVETRIEVRPCGOPYSSLKKGKSKTKSPBPV 296
DB 206 SACSCTGCMGISTRTVNDNPECRILVETRIEVRPCGOPYSSLKKGKSKTKSPBPV 265
QY 297 RFTYAGCLSVKRYKPGYCGSVDCRCCTPOLTRIVYKMFREDGETFSKNMMQOSCKN 356
DB 266 KFEISGCTSVKRYKPGYCGSVDCRCCTPOLTRIVYKMFREDGETFSKNMMQOSCKN 325
QY 357 YNCPHANEAFPEY--RLFNDI 376
DB 326 YNCPGNDIFESILYRKMYGDM 347

RESULT 14
US-10-053-753-6
Sequence 6, Application US/10053753
Publication No. US20020150986A1
GENERAL INFORMATION:
APPLICANT: Iau, Lester F.
TITLE OF INVENTION: Extracellular Matrix Signalling Molecules
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Marshall, O'Toole, Garstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

```

; APPLICATION NUMBER: US/10/053,753
; FILING DATE: 22-Jan-2002
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
;   NAME: Clough, David W.
;   REGISTRATION NUMBER: 36,107
;   REFERENCE/DOCKET NUMBER: 28758/33766
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 312/474-6300
;   TELEFAX: 312/474-0448
;   TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 6:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 348 amino acids
;     TYPE: amino acid
;     STRANDEDNESS: single
;     TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
;   NAME/KEY: misc-feature
;   OTHER INFORMATION: "Fisp12 amino acid sequence"
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-10-053-753-6

```

Query Match 44.8%; Score 948.5; DB 14; Length 348;

Best Local Similarity 46.3%; Pred. No. 2.5e-73;

Matches 177; Conservative 58; Mismatches 106; Indels 41; Gaps 8;

```

QY 1 MSSRIARALAIYVTLIHL-TRIAL-STCPAACHPLE-APKCAPGVGLVRDGGCCCKYCA 57
   | : : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
Db 1 MASVAGPISIALVILALCTRPATGODCSAOCOAEPAPHPAGVSLVDGCGCCRYCA 60
QY 58 KOLNEDSKTOPCDHTKGLKCNFGASTALKGICRAQSEGRPCVNSRIYONGESPQNC 117
   | | | | : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
Db 61 KQGLGELCTERDPCDDPKHGLFCDFGSPANKRIGVCTAK-DGAPCVFGSVYRSGESPOSSC 119
QY 118 QHOCTCIDGAVGCIPLCPQELSLPNLGCNRPRLVKTGGCCCEWVDEDSINDPMEDQG 177
   : : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
Db 120 KYQCTCLDAGVGVPLCSMDVRLPSPDCFPFRVKLPKCKCKEWVDEP----- 168
QY 178 LQKELGFDASEVELTRNNELIAGKGRSLKRL-PVFGMEPRILYNPLQGGKCIYQTSM 236
   | | | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
Db 169 -----KDRTAVGPAALAYRLIEDTFGDPPTM-----RANCLVQTTM 205
QY 237 SOCSKTGIGTSTRVTNDNPECRLYKETRICEVRCGQPVYSSLKKGKCSKTKSPQV 296
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 206 SACSKTGCMGISTRTVNDNTFCLEKOSRLCMVRPCEADLEENIKKGCIRTPKIAKPV 265
QY 297 RFTYAGCLSVKKRPKYCGSCVDGRCTPQLRTVAMRRCDEGEFFSKNVMAMIOSCKN 356
   : : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
Db 266 KEELSGCTSVKTYRAKFCGCTDGRCTPHRTTTLVEFRCPDGEIMKKNMFIKTQACH 325
QY 357 YNCPHANEAPFPY--RLFNDI 376
   | | | | : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
Db 326 YNCPGNDLFEESLYYRKMGDM 347

```

RESULT 15

US-10-245-977-8

; Sequence 8, Application US/10245977

; Publication No. US20030113816A1

; GENERAL INFORMATION:

; APPLICANT: Weitz, Stephen L

; APPLICANT: Usinger, William R

; TITLE OF INVENTION: METHODS OF ASSAYING CONNECTIVE TISSUE GROWTH FACTOR

; FILE REFERENCE: FP0812 US

; CURRENT APPLICATION NUMBER: US/10/245,977

; CURRENT FILING DATE: 2002-09-18

; PRIOR APPLICATION NUMBER: US 60/323,305

; PRIOR FILING DATE: 2001-09-18

; NUMBER OF SEQ ID NOS: 8

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 8

```

; LENGTH: 348
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-245-977-8

```

Query Match 44.8%; Score 948.5; DB 15; Length 348;

Best Local Similarity 46.3%; Pred. No. 2.5e-73;

Matches 177; Conservative 58; Mismatches 106; Indels 41; Gaps 8;

```

QY 1 MSSRIARALAIYVTLIHL-TRIAL-STCPAACHPLE-APKCAPGVGLVRDGGCCCKYCA 57
   | : : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
Db 1 MASVAGPISIALVILALCTRPATGODCSAOCOAEPAPHPAGVSLVDGCGCCRYCA 60
QY 58 KOLNEDSKTOPCDHTKGLKCNFGASTALKGICRAQSEGRPCVNSRIYONGESPQNC 117
   | | | | : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
Db 61 KQGLGELCTERDPCDDPKHGLFCDFGSPANKRIGVCTAK-DGAPCVFGSVYRSGESPOSSC 119
QY 118 QHOCTCIDGAVGCIPLCPQELSLPNLGCNRPRLVKTGGCCCEWVDEDSINDPMEDQG 177
   : : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
Db 120 KYQCTCLDAGVGVPLCSMDVRLPSPDCFPFRVKLPKCKCKEWVDEP----- 168
QY 178 LQKELGFDASEVELTRNNELIAGKGRSLKRL-PVFGMEPRILYNPLQGGKCIYQTSM 236
   | | | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
Db 169 -----KDRTAVGPAALAYRLIEDTFGDPPTM-----RANCLVQTTM 205
QY 237 SOCSKTGIGTSTRVTNDNPECRLYKETRICEVRCGQPVYSSLKKGKCSKTKSPQV 296
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 206 SACSKTGCMGISTRTVNDNTFCLEKOSRLCMVRPCEADLEENIKKGCIRTPKIAKPV 265
QY 297 RFTYAGCLSVKKRPKYCGSCVDGRCTPQLRTVAMRRCDEGEFFSKNVMAMIOSCKN 356
   : : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
Db 266 KEELSGCTSVKTYRAKFCGCTDGRCTPHRTTTLVEFRCPDGEIMKKNMFIKTQACH 325
QY 357 YNCPHANEAPFPY--RLFNDI 376
   | | | | : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
Db 326 YNCPGNDLFEESLYYRKMGDM 347

```

Search completed: August 5, 2003, 14:18:51
Job time : 35.5908 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 5, 2003, 14:03:01 ; Search time 15.5408 Seconds

(without alignments) updates/sec
2357.684 Million cell

Title: US-09-495-448A-4

Perfect score: 2116
Sequence: 1 MSSRIARALAVLTLLHLTR.....ANEAFPRYLNDIHKFRD 381

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1929	91.2	379	2	A35669 gene CYR61 protein
2	1695	80.1	375	2	A41428 CER-10 protein pre
3	952.5	45.0	348	2	A40578 beta IG-M2 protein
4	947	44.8	349	2	A40551 connective tissue
5	851.5	40.2	351	2	S20078 NOV protein - chic
6	827.5	39.1	357	2	138069 gene novh protein
7	174	8.2	1111	2	T26972 hypothetical prote
8	173.5	8.2	1620	2	T27283 hypothetical prote
9	157	7.4	1700	2	S08167 Baldini ring 3 pr
10	153	7.2	1574	2	T13954 MEGF6 protein - ra
11	152	7.2	5376	2	T42215 zonadhesin - mouse
12	150.5	7.1	1178	1	A39804 Thrombospondin pre
13	150	7.1	837	2	A42112 mucin-like peptide
14	147.5	7.0	1170	2	A53612 lam1n Blk chain
15	146	6.9	1964	2	T09059 notch4 - mouse
16	145	6.9	601	2	T22025 hypothetical prote
17	145	6.9	601	2	D89711 protein FA0E10.4 l
18	141.5	6.7	13288	2	T03099 mucin, submaxillar
19	141	6.7	1042	2	A57534 mucin 5AC (clone L
20	141	6.7	3020	2	A43932 mucin 2 precursor,
21	140.5	6.6	1023	2	T42626 secreted leucine-r
22	140.5	6.6	1034	2	JC5598 mucin - rat
23	140	6.6	4135	2	T42629 tenascin-X - bovin
24	139.5	6.6	1101	2	T16840 transmembrane prot
25	139	6.5	2437	2	S42620 notch protein - fr
26	138.5	6.5	2703	1	A24420 mucin MUC5B, trach
27	138	6.5	1056	2	A53767 slit protein 1 pre
28	137.5	6.5	1480	2	A36665 notch protein homo
29	137	6.5	2531	2	S18188

30	135	6.4	1168	2	I56985 kalinin B1 - mouse
31	135	6.4	1170	2	A40558 thrombospondin 1 p
32	134.5	6.4	330	2	T25169 hypothetical prote
33	134.5	6.4	4006	2	T09070 probable tenascin
34	134	6.3	1959	1	AGRT agrin - rat
35	134	6.3	3106	1	S53868 lam1n alpha-2 ch
36	134	6.3	3672	2	T23433 hypothetical prote
37	134	6.3	3704	2	T37316 probable lam1n a
38	133.5	6.3	1955	1	AGCH agrin precursor -
39	133	6.3	2813	1	VWU von Willebrand fac
40	132.5	6.3	2555	2	A40043 notch protein homo
41	132	6.2	251	2	A55035 cysteine-rich pre
42	132	6.2	1469	2	B36665 slit protein 2 pre
43	132	6.2	1531	2	T42218 slit-l protein hom
44	132	6.2	2823	2	T23064 hypothetical prote
45	132	6.2	2823	2	F87908 protein T22A3.8 l1

ALIGNMENTS

RESULT 1

A35669

gene CYR61 protein precursor - mouse

C:Species: Mus musculus (house mouse)

C>Date: 28-Sep-1990 #sequence_revision 18-Nov-1992 #text_change 05-Nov-1999

C:Accession: A35669; I48319; S16446

R:O'Brien, T.P.; Yang, G.P.; Sanders, L.; Lau, L.F.

Mol. Cell. Biol. 10, 3569-3577, 1990

A>Title: Expression of cyr61, a growth factor-inducible immediate-early gene.

A:Reference number: A35669; MUID:90287146; PMID:2355916

A:Accession: A35669

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-379 <OAB>

A:Cross-references: GB:M32490; NID:9192909; PIDN:AAA37512.1; PID:9309206

A>Note: the authors translated the codon GAT for residue 337 as Gln

R:Latinkic, B.V.; O'Brien, T.P.; Lau, L.F.

Nucleic Acids Res. 19, 3261-3267, 1991

A>Title: Promoter function and structure of the growth factor-inducible immediate ear

A:Reference number: I48319; MUID:91288203; PMID:2062642

A:Accession: I48319

A>Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-379 <RES>

A:Cross-references: EMBL:X56790; NID:950632; PIDN:CAA40109.1; PID:950633

A>Note: the authors did not translate the codon for residue 108

A:Note: the authors translated the codon GAT for residue 337 as Gln

C:Genetics:

A:Gene: CYR61

A:Introns: 21/3; 93/1; 208/1; 279/3

C:Superfamily: von Willebrand factor type C repeat homology <WVC>

F:199-166/Domain: von Willebrand factor type C repeat homology <WVC>

Query Match 91.2% Score 1929; DB 2; Length 379; Best local similarity 90.9% Pred No. 1.2e-134; Mismatches 19; Indels 6; Gaps 2;

QY	1	MSSRIARALAVLTLLHLTRALSTCPAACHCPLEAPCAPGVGLVRDGGCCRYCARQL	60
DB	1	MSSRIARALAVLTLLHLTRALSTCPAACHCPLEAPCAPGVGLVRDGGCCRYCARQL	60
QY	61	NEDCKTQPCDHTKLECNFGASSALAGICRAQSEGRCEYNSITYNGESFQNCQH	120
DB	61	NEDCKTQPCDHTKLECNFGASSALAGICRAQSEGRCEYNSITYNGESFQNCQH	120
QY	121	CTCIDGAVGCIPLCPQELSLPILGCPNRLVAVTGCCEEWYCDSDSIKDPEDDGLG	180
DB	121	CTCIDGAVGCIPLCPQELSLPILGCPNRLVAVTGCCEEWYCDSDSIKDPEDDGLG	180
QY	181	KELGFDASEVELTRNNELIAVGKGRSLRLPVFGMEPRILYNPL--QGOKIVQTTSMQ	238
DB	179	--LGIDASEVELTRNNELIAVGKGRSLRLPVFGMEPRILYNPLAHHQOKIVQTTSMQ	236

QY	239	CSATCGGISTRTVNTNNPECLVYKERIEVBPCCOPYASSLKGKKCKSKTKKPEPRF	299
		:	
Db	237	CSASCGETISTRTVNTNNPECLVYKERIEVBPCCOPYASSLKGKKCKSKTKKPEPRF	296
QY	299	TYAGCLSVKKYRBYKCGSCVDRCCKTPOLTRPVKKRPFCEDETFPSKVMVMIQSCCKYN	358
Db	297	TYAGCSSVKYRBYKCGSCVDRCCKTPOLTRPVKKRPFCEDETFPSKVMVMIQSCCKYN	356
QY	359	CPHANEAAPFYRLFNDIHKFRD	381
		:	
Db	357	CPHANEAAPFYRLFNDIHKFRD	379

RESULT 2

CEF-10 protein precursor - chicken
 C:Species: Gallus gallus (chicken)
 C:Date: 03-Apr-1992 #sequence_revision 03-Apr-1992 #text_change 21-Jul-2000
 C:Accession: A41428
 R:Simmons, D.L.; Levy, D.B.; Yannoni, Y.; Erikson, R.L.
 Proc. Natl. Acad. Sci. U.S.A. 86, 1178-1182, 1989
 A:Title: Identification of a phorbol ester-repressible v-src-inducible gene
 A:Reference number: A41428; M01D:89145206; PMID:2537491
 A:Accession: A41428
 A:Status: Preliminary
 A:Molecule type: mRNA
 A:Residues: 1-375 <SIM>
 A:Cross-references: GI:J04496; NID:g21i1435; PIDN:AAA48661.1; PID:g21i1436

Query Match	80.1%;	Score 1695;	DB 2;	Length 375;
Best Local Similarity	81.0%;	Pred. No. 2e-117;		
Matches 311;	Conservative 20;	Mismatches 41;	Indels 12;	Gaps 7

QY	I	MSRIRALALAYVTLHLHLRAL - SICPAACHCIPLEAKRCAPGVGLVBDGGCCCKYAKQ	59
Db	1	MGASAGAR - ALAALLCLARLALGSCPCAPVCCPPAAAPQCAPGVGLVPDGGCCCKYAKQ	59
QY	60	LNEDCSKTQPCDHTKTELECNFCGASSPALKGICRAOSEGRPCPYNSRIYONCESFPQCOH	119
Db	60	LNEDCSKTQPCDHTKTELECNFCGASPAATNGICRAOSEGRPCPYNSKIYONCESQPPCKH	119
QY	120	QCTCIDGAVNCIPLCPQELSPENLPGCPNBRLYKVYQCCCEWYCDDEDSIKDPMEDODGLL	179
Db	120	QCTCIDGAVNCIPLCPQELSPENLPGCPNBRLYKVYQCCCEWYCDDES - KALDELBSGF	177
QY	180	GRELGDASVELTRNNELIANVGKGRSLKRLPFVGMER - RILNPNLOGQKCIYQTTWS	237
Db	178	SKREGLDASBEGELTRNNELIALTVK - GLKMLPFVFSBPQSAFENP - - - KCIYQTTWS	232
QY	238	QCSKTGCTGSTFVTNDNPECRILVETRICERVPCGQVYSSLKGGKCSKTKSPSPVR	297
Db	233	QCSKTGCTGSTFVTINDNPDCKLIETRICERVPCGQVSYSLKGGKCTYTKKSPSPVR	292
QY	298	FTYAGCLSVKKYKPKYCGSCVDRCCTPOLFTFYKMRPCEBDGETFSKNVMIMIOSCRNY	357
Db	293	FTYAGCSSVKKYKPKYCGSCVDRCCTPOQFTFYKIRPCDDGETFTSKVMIMIOSCRNY	352
QY	358	NCPHANEAAPFPRFLNDIHKFRD	381
Db	353	NCPHANEAAPFPRFLVNDIHKFRD	375

RESULT 3

beta IG-M2 protein precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 06-Mar-1992 #sequence_revision 06-Mar-1992 #text_change 01-Dec-2000
C:Accession: A40578; A53228
R:Brummer, A.; Chinn, J.; Neubauer, M.; Purchio, A.F.
DNA Cell Biol. 10, 293-300, 1991
A:Title: Identification of a gene family regulated by transforming growth factor-beta
#:Reference number: A40578; MUID:Y1292699; PMID:2029337

A:Accession: A40578
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-348

A:Cross-references: GB:M60263; NID:g201945; PIDN:AAA73135.1; PID:g201946
R:Rysecq, R.P.; Macdonald-Bravo, H.; Mattei, M.G.; Bravo, R.
Cell Growth Differ. 2, 225-233, 1991
A:Title: Structure, mapping, and expression of flsp-12, a growth factor-inducible gene
A:Reference number: A53228; MUID:91363290; PMID:1888668
A:Accession: A53228
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-160, 'K', 162-348 <RYS>
A:Cross-references: GB:M70641; NID:g193313; PIDN:AAA37627.1; PID:g193314
C:Genetics:
A:Gene: flsp-12

Query Match	45.0%;	Score 952.5;	DB 2;	Length 348;
Best Local Similarity	46.6%;	Pred. No. 7,6e-53;		
Matches 178;	Conservative 57%;	Mismatches 106;	Indels 41;	Gaps 8

```

OY      1  MSSRIKRALALVYTLHL-TRLAL-STPPAACHCLE-APKCAPGVGLVPGCCGCAVCA  57
Db      1  MASVAGPISLALVYTLALCTLRATGQDCAAGCGCAAAAPHCAPAGVSLVLDGCCCAVCA  60

OY      58  KQINEDCSKTQPCDHTKGLGECNFGASSALAKGICRAOSEGRPCREYNSTIYONGESFOPNC  117
Db      61  KQIGELICTERDPDHPKGLGFCDFGSPANKRIQVCTAK-DGAPCVAGSGVYSRSGESFOSSC  119

OY      118  OHQCTCIDGAVGCIPLCPQELSLPNCGPNRPVLVYTGCCCEBWCCDEDSIKDPMEDQG  177
Db      120  KYQCTCIDGAVGCVPLCSMDVRLPSPDCCFPFRXVALPKKCCBEMWCDP-----  168

OY      178  LIGKELGFDASEVELTERNNELIAYAGSGRSIKRL-VFGEEMERILYVLOGOKCIQVTTSM  236
Db      169  -----KDRITAVGPBALAAYLLEDTFGPDPTMM-----RANCLVOTTEW  205

OY      237  SQCSKTCGTGISTRYVNDNPNPECRILYKETRICEVRPCGQPVYSLSLKGGKCSKTKKSPDY  296
Db      206  SASCKTCGMSISTRYVNDNTNPFCLRKQSRSLCYNVRCEADLLENIKKKGKCIPTPKIAKP  265

OY      297  RFTYAGCLSYKKTRPYKCGSVYDGKCTPQLTRLYKAMFREDDEFTSKNVMIIQSKCN  356
Db      266  KPELSGCTSVKTYRAAFKCGCVYDGCCTPRHTTTLVPEFKCPDEGIMKKMMFIKTCACH  325

OY      357  YNCPHANEAAPFY--RLFNDI  376
Db      326  YNCPGNDLPESLIYKRMIGDM  347

```

RESULT 4

connective tissue growth factor - human
A:Species: Homo sapiens (man)
C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 21-Jul-2000
C:Accession: A40551; S44205
R:Bradham, D.M.; Igarashi, A.; Potter, R.L.; Grotendorst, G.R.
J. Cell Biol. 114, 1285-1294, 1991
A:Title: Connective tissue growth factor: a cysteine-rich mitogen secreted by human v
A:Reference number: A40551; MUID:91373462; PMID:1654338
A:Accession: A40551
A:Molecule type: mRNA
A:Residues: 1-349 <BBA>
A:Cross-references: GB:M92934; GB:M36965; GB:S56201; NID:g180923; PIDN:AA91279.1; PI
R:Oemar, B.S.; Wenner, A.; Yang, Z.; Garnier, J.M.; Gentz, R.; Luescher, T.F.
submitted to the EMBL Data Library, April 1994
A:Description: Differential cloning and expression of human connective tissue growth
A:Reference number: S44205
A:Accession: S44205
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-349 <DEM>
A:Cross-references: EMBL:X78947; NID:g474933; PID:g474934

Query Match	44.8%;	Score 947;	DB 2;	Length 349;
Best Local Similarity	46.68;	Pred. No. 1.9e-62;		
Matches 174; Conservative	58;	Mismatches 101;	Indels 40;	Gaps 8;

```

QY      9  IAVVTLHLTRLAL-STCPAACHIPLE-APKCAPGVGLVADGGCCCKVAKOLNEDCSK 66
Db      11 VAFVVLALCSRPVAVGONCSGPCRCBPDEPARPCAGVSLVLDGGCCRVAKOLGELCTE 70
QY      67 TOPCDHTKGLCEJNAGASSTALKICIGRAOSEGRPEVNSRIYONCESFOPNCOHOCTCIDG 126
Db      71 RDPCDDPHKGLFCDPDGSPANKRIGVCTAK-DGAPCIFEGTYIRSESESQSKRYOCTCIDG 129
QY      127 AVGCIPICPELSPNLNIGCPNPRLVKVTGOCCEEWCDSDSIKDPMEDODGILKELGFD 186
Db      130 AVCGMPICSDMDVRLPSPDCPFPRVRKLPKGCCEEWCDSE-----KDQ----- 172
QY      187 ASEVELTRNNELLAVGGRSLKRL-PVFGMEPRLLVPLQOGKCIIVOTTSMSQCSKTGCT 245
Db      173 -----TVGPAALAYRLDFTFGDPPTMI-----RANLIVOTTESSAKSCTGGM 215
QY      246 GISTRVTNDNPECKLVKETRICEVRPCGQPYSSLKKGKCSKTKKDEPRFRTYACLS 305
Db      216 GISTRVTNDNASCHLEKOSRCLCMVRPEADLEENIKKGKKCIRPKTSKPKFELSGCTS 275
QY      306 VKKIRPYKCGSCVDGRCTPOLFTVTMARCEDEGEFFSKNVMMIOCSKCYNCPHANEVA 365
Db      276 MKTYBRAFCVCTDGRCCCTPHRTTLLVPERKCPDGEVKKNNMFIKTCACAYNCPGNDI 335
QY      366 AFPEFY--RLENDI 376
Db      336 FESILYRKMYGM 348

```

```

RESULT 5
S20078
NOV proteIn - chicken
C:Species: Gallus gallus (chicken)
C:Date: 19-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 21-Jul-2000
C:Accession: S20078
R:Joblot, V.; Martinierle, C.; Dambrine, G.; Plassart, G.; Brianc, M.; Crochet, J.; Pertin
Mol. Cell. Biol. 12, 10-21, 1992
A:Title: Proviral rearrangements and overexpression of a new cellular gene (nov) in myel
A:Reference number: S20078; MUID:92107157; PMID:1309586
A:Accession: S20078
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-351 <JOL>
A:Cross-references: EMBL:X59284; NID:963702; PIDN:CAA1975.1; PID:963703
C:Genetics:
C:Gene: NOV

```

Query Match	40.28	Score 851.5	DB 2	Length 351
Best Local Similarity	44.08	Pred. No. 2.1e-55		
Matches 164	Conservative 44	Mismatches 104	Indels 61	Gaps 8
QY	11	LVTLLHLTRLALS---	TCPAAC--HCPLEAPKADAVGLVVRDGCGCCKVCYAKQLNEDC	64
		::	::	
Db	12	LLTLTLTLRRCFEVSGREACAPRCQGRCPAPBPACGAVAVLDGSCCCCLVCARQRESC	71	
		::	::	
QY	65	SKTQPCDHTGRLCENFGASSTALKGTCRASEGSPCYNSRITYONGESPQPNCHOCTCI	124	
		::	::	
Db	72	SPLLPCDESGLVCDKRPEDGGAGTGMV--LEGDNCVFDMGKIRNGEFTFQSCYQCTCR	130	
		::	::	
QY	125	DGAVGCIPLCPQBELSLPNIGCNPBRILVYKTCGCEEWACDEDSIKDEMODGILLG----	180	
		::	::	
Db	131	DGIGICLPRCNLTLLLPGRDCCPEPRKRIEYGECECKWVC-----DPRDEV--LLGSGFAM	182	
		::	::	
QY	181	-----KELGFDASEVELTRNNELLIAVGKRSILKRLVFMERPRILNLPLOGQKCIYOTT	234	
		::	::	
Db	183	AATRGQENTLGIIDVSD-----SSANCIQOTT	207	
		::	::	
QY	235	SWGQCSKTGCTGISTRAYTNDNPBCRLVKETRICEVRPCGGPYVSLIKGKRSCTKRSPE	294	
		::	::	

[illegible]

RESULT 6
138069
gene novh protein - human
CISpecies: Homo sapiens (man)
C.Date: 17-May-1996 #sequence_revision 17-May-1996 #text_change 05-Nov-1999
C.Accession: I38069
R.Martinerie, C.; Huff, V.; Joubert, I.; Badzioch, M.; Saunders, G.; Strong, L.; Petb
Oncogene 9, 2729-2732, 1994
A>Title: Structural analysis of the human nov proto-oncogene and expression in Wilms
A.Reference number: I38069; MUID:94336229; PMID:7520150
A.Accession: I38069
A:status: preliminary; translated from GR/EMBL/DBD
A:Molecule type: DNA
A:Residues: 1-357 <RES>
A:Cross-references: EMBL:X78351; NID:9587422; PIDN:CAA55146.1; PID:9825696
C:Genetics:
A:Gene: novh
A:introns: 28/3; 104/1; 188/1; 259/3
C:Superfamily: thrombospondin type 1 repeat homology
F703-250/Domain: thrombospondin type 1 repeat homology <THR1>

Query Match	Similarity	39.1%	Score 827.5	DB 2:	Length 357:
Best Local	Similarity	43.1%	Pred. No. 1,2e-53:		
Matches	157;	Conservative: 148;	Mismatches 120;	Indels 39;	Gaps 7
Db	7	RALALVTLHLHTLRLALST	-CPAAC	-HCPLEAPKCAPGVGLVBDGCGCCVYCAKOLNE	62
QY	14	QCLCTFLTLLHLLGLGVAATRCPPCCPKCPATPTPTCAAGVAVLBDGCECCVLCARQGE			73
QY	63	DCSKTQPCPDHMKGLECNFGASSTALKGICRQASERPCPEYNSRITONGESSFQPCNOHQCT			122
Db	74	SCSDLEPDESSGLCLCDRSADPSNOTGICTA	-VEBDCNCFDSDVITYRSKGKFPQSKFOOT		132
QY	123	CIDGAVGICPLCPQPLSLPNIKGCPNRIPLVYKVGCCSEEWCDSDSDIKDPMEOQDGLTKE			182
Db	133	CRDQIGCVPCQGLDVLPLPEPNCAPFRVYEVGCECKWIGCPD	-----EDSLGLGT		185
QY	183	LGFDASEVELTRNNELIIVGKGRSLKRLPVGMEIRILYNPLQOGKCIYOTTSMGOSCKT			242
Db	186	LAAYRPEATL	-----GVEVSD	-----SSVNCIEOTLEWTMACSKS	219
QY	243	CGTGISTFVTNDNPECRVLKTRICEVAPCGO	-PYSSILKKGKCKSKTKSPPEVPRTYA		301
Db	220	CGMGESTVATYNNRQCEMLKOTRLICMVRPEDEPQOPIDKKGKKCLRTKSKLAIHLDFK			279
QY	302	GCLSVKRYKPKYCGSCVVDGCCCTPOLRTVTKMRFCEDEGTFSSKNVMIQSCKNCVNCBH			361
Db	280	NCTSLHTYKPRFCGVCSDGRCCTPHNTKTIQAEFQCSGQILVKKRPMVVIIGTCTCHNCBK			339
QY	362	ANEA	365		
Db	340	NNEA	343		

RESULT 7
T26972
hypothetical protein Y47H9C.4 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 17-Mar-2000
C:Accession: T26972
R:Harris, B.

submitted to the EMBL Data Library, October 1998

A:Reference number: Z20293

A:Accession: T26972

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-1111 <MIL>

A:Cross-references: EMBL:AL033657; PIDN:CAA21739.1; GSPDB:GNO0019; CESP:Y47H9C.4

A:Experimental source: clone Y47H9C

C:Genetics:

A:Gene: CESP:Y47H9C.4

A:Map position: 1

A:Insertions: 50/2; 84/2; 150/1; 238/3; 342/3; 797/1; 851/1; 947/2; 1017/1; 1083/1

C:Superfamily: unassigned ankyrin repeat proteins; ankyrin repeat homology; EGF homology

Query Match

Best local similarity 19.7%; Pred. No. 4.3e-05;

Matches 84; Conservative 33; Mismatches 161; Indels 148; Gaps 17;

```

QY 26 CPAACHCPLP-APKAPGVGLVR-----DGGCGKVCAR-QLNEDSKTPQCDHTKGLC 78
DB 380 CSKTCCTVBNELTMCAPNTGFCRCRKGFGDNCCLASCKDSYGPNCCKQAMCDMNHASEC 439
QY 79 NNGASTALKGICRAQSEGRPCYNSRIYONGESPQPCQHOCTC-----IDGA----- 127
DB 440 NPETGSC---VCKPRTGKNCSEPCPL---DEYPCNAHQCCQCMORGVCDCGADGCKQ 491
QY 128 -----VGCIPPL-----CPQELSLPN--LGCP-- 146
DB 492 CQKGTWGHCEHHCAPDTFGANCKRCKRCKGICGCPITGECGCPAGLOGANDIDGCPES 551
QY 147 -----NPLRVYTGQCCCEWVCDSDSIKDPMEDDGLLGLKELGPDASEVELT 193
DB 552 SYGPGCKLHCKVCWNGCDKDETEGC---TC-----QPFSGSCSTTCSKGYG 596
QY 194 RNNELIYAVKGRSLKRLPFQGMERILYNPLOGQKCIYQTSQSCSKTCGTISRNVN 253
DB 597 ESECL-----SCPCSDASCKTGKCLPLGTRGV 627
QY 254 DNEPCRLVETRICVPRPCGPVYSILKCKKCKSKTRKSPPEVFTY-----AGCLS 305
DB 628 CDQKCDPNTFGLCGFTVTPSPCASIDPKNGVCLSPDPSSGJHCEHNPAGSYGDCQ 667
QY 306 VVKYRKYGSCYDGCCTPQLTRIV-----KRFRCDEGETSKNVMITQSKCNT 357
DB 688 V-----C-SCADHGCDPPTGECICEPBGHGKTCSEKCPDKYGYGALDCKPCASGS 739
QY 358 NCPHAN 363
DB 740 TCDHIN 745

```

RESULT 8

T27283 hypothetical protein Y64G10A.f - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T27283

R:Almscough, R.

submitted to the EMBL Data Library, September 1999

A:Reference number: Z20336

A:Accession: T27283

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-1620 <MIL>

A:Cross-references: EMBL:AL110498; NID:el542303; PIDN:CA854471.1; CESP:Y64G10A.f

A:Experimental source: clone Y64G10A

C:Genetics:

A:Gene: CESP:Y64G10A.f

A:Insertions: 77/1; 116/1; 198/1; 282/1; 365/1; 425/1; 466/1; 548/1; 559/1; 601/1; 625/1;

Query Match

Best local similarity 24.2%; Score 173.5; DB 2; Length 1620;

Matches 92; Conservative 33; Mismatches 146; Indels 109; Gaps 21;

```

QY 26 CPAACHCPLPAP-----KCAPG-----VGLVRDGGCCCKVCARQALNEDCSK 66
DB 1052 CKGICSCQNGATDSDVTSGSECRPGWRKAKCDRCPDG--RFEGGNALICDCTTTNDTSM 1109
QY 67 TDP---CDHTKLECNFGASSTALKGICRAQSEGRPCYNSRIYONGESPQPCQHOCT 122
DB 1110 YNPFVARCHVHTG-ECR-----CPAGTGDCCQTSGLGHNGE---GCHRSQ 1153
QY 123 CIDGAVGCIPLCPQELSLPNLNGCPNRLVKTGQCEWVCDSDSIKDPMEDDGLLGLKE 182
DB 1154 CSNGA-----SCRVYTF--CDPSPGFMNCSECEPGLMGSKCKHCLMHGECNKE 1206
QY 183 LGPDASEVELTRNNELIYAVKGRSLKRLPFQGMERILYNPLOGQKCIYQ-----TSM 236
DB 1207 NG-----DCEIDGWTGSLDPFGQFGNCAQRCKNKGASCDRRTGCECLPGW 1256
QY 237 S--QCSKTGTGTISTVTNDNPECRVLYETRIEVRPCGPVYSSLK-----KGRCKST 289
DB 1257 SGHECHKSCVSG-----HYGARC---ETCEBENGALCDPISGHSCQPGWGRKCKNR 1307
QY 290 KSPPEVRYFTYACLSVKKYRKYCG-----SC--VDGRCTPQLTRVYKMRFC 338
DB 1308 -----CL--KGFGRHCSQSCRANKSCDHI SGRCQPKYAGHSTELCP 1352
QY 339 DGETFSKNVMITQSKCNTN 358
DB 1353 DG-TFEGSCS--QKDCGEM 1369

```

RESULT 9

S08167 Balbiani ring 3 protein - midge (Chironomus tentans)

C:Species: Chironomus tentans

C:Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 21-Jul-2000

C:Accession: S08167

R:Paulsson, G.; Lendahl, U.; Galli, J.; Ericsson, C.; Wieslander, L.

J. Mol. Biol. 211, 331-349, 1990

A:Title: The balbiani ring 3 gene in Chironomus tentans has a diverged repetitive str

A:Reference number: S08167; MID:90172404; PMID:1689777

A:Accession: S08167

A:Status: not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-1700 <PAU>

A:Cross-references: GB:X52263; NID:97057; PIDN:CAA36506.1; PID:97058

C:Genetics:

A:Gene: BR3

A:Map position: 4

C:Superfamily: unassigned Balbiani ring proteins

Query Match

Best local similarity 7.4%; Score 157; DB 2; Length 1700;

Matches 88; Conservative 47; Mismatches 170; Indels 116; Gaps 21;

```

QY 30 CHCPLPAPK--CAPGYGLVRDGGC--CKYCAKOLNEDSKTYQPCDHTKGLC-----NNG 81
DB 221 CICTPTEPAGGCSAPLKWDDKSCACAPKMEKKEKCVESGKIMVNTCEGCAQLNCP 280
QY 82 ASSTALKGICRAQ--SEGRPCYNSRIYONGESPQPCQHOCTCI----- 124
DB 281 DNKKANKETGQCCCKEYKCC-----NGQVF--CKSSCSVCYGGGDKDKTCTAPQY 330
QY 125 DGAVGCIPLCPQELSLPNLNGCPNRLVKTGQCEWVCDSDSIKDPMEDDGLLGLKE 184
DB 331 DG-VACSCSPVYVMQKPADGCPRP-----QKWDRECHCECPVR-HDCKNKGKVM 378
QY 185 PDASEVELTRNNELIYAVGGR-----SLKRLPFQGMERILYNPLOGQKCIYQTSM 236
DB 379 ETIQGQICPRDAPVCTAGKERGCESECKCINNEPREGAKPLVWNE-NTCKCVCPADK- 436
QY 237 SQSKTCGTGISTRTVNDNDEC-----RLVETRIEVRP-----GOAPVYSS 280
DB 437 QMSPGGCGSGKSNKLTCCQCECDQASKGLKRMNADTKCKECQPGMPPEGCGKQWISD 496

```

```

OY 281 KKGKCKSKT-----KSPPEVPTVAGLSVKKYR--- 311
      |||
Db 497 KKCCESSPTTQAPQILDINTCECKCPVMLAOKCKSPKOWTDSKCLCECSTTPATC 556
OY 312 ----KYCG---SCV---DGRCCPTPOLTFRVKKRFRCEDEET-FSKVMMIIOCKNYNC 359
      |||
Db 557 EKGOTMCGEACGCLCPGDCNCKNKKFDPKSCCECKKNPTCTSPVMMADDEC--KC 614
OY 360 P 360
      |
Db 615 P 615

RESULT 10
T13954
MEGF6 protein - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jul-2000
C:Accession: T13954
R:Nakayama, M., Nakajima, D.; Nagase, T.; Nomura, N.; Seki, N.; Ohara, O.
Genomics 31, 27-34, 1998
A:Title: Identification of high-molecular-weight proteins with multiple EGF-like motifs
A:Reference number: Z14126; MUID:98360089; PMID:9693030
A:Accession: T13954
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1574 <NA>
A:Cross-references: EMBL:AB011532; NID:g3449293; PIDN:BAA32462.1; PID:g3449294
A:Experimental source: strain Sprague-Dawley; brain
C:Genetics:
A:Gene: MEGF6

Query Match 7.28; Score 153; DB 2; Length 1574;
Best Local Similarity 21.94; Pred. No. 0.0021;
Matches 97; Conservative 36; Mismatches 133; Indels 176; Gaps 28;

OY 26 CPAACHCP-----LEAPKCAPV--GLVRDGC----- 50
      |||
Db 568 CSSPCTQNGCTDPPVLAGACRCPGVSAGHCEGCKPGYTKHCKKCHCANRGCHRLY 627
OY 51 GCCKVCAPKOL-----NEDCKTOPCDHTKGLCNFGASSTALKICR 92
      |||
Db 628 GAC-ICDPGLYGRFCHLACPPMAFGPCSDCLCEQ--SHTRSCNPKDSCS-----CK 678
OY 93 AQBEGRPCEYNSRIYONGESFQPCOHCCTCIDGANGCIDL-----CPQ 136
      |||
Db 679 AGFQGERCAEC---ESG--FFGPGCRHRCYCPG-VACDPVSGECRTQCPGYOGEDCGQ 733
OY 137 ELSLPNG-----CPNRLVKNYQC--CEMWCDSDSIKPMEDQDLKELGFDA 187
      |||
Db 734 ECPVGTGGVNCSSCSVGPACHRVTEGCLCPKGTGEDGACDPESRWELGCOEI--- 789
OY 188 SEVELFRNNELIIVAGKRSILKRLPVFGMEPRILYNPLQOKCIYQTSMSQCKTC--- 243
      |||
Db 790 -----CPACRCHASC-----NPEGTGCLCLPGVVSRCQDPTGACW 825
OY 244 -GGIGISTR--VTNDNPECLRVKETRICEVRPCQPVYSILKKGKCKSKTKSPPEVPTY 300
      |||
Db 826 YGTGCGCIRACACANDG-HCDPTTGRSC-----APGWTGLSCIRACDSGHMGDP----- 872
OY 301 AGLSVKKYRKYC---GSC--VDGR-CT-----POLRLTYKMRP--RCDDGEFFSK 345
      |||
Db 873 --CI-----HPCMSAGHGKCDVAVSGCLCEAGYEGRCRCHSCQGYGSPCE----- 918
OY 346 NVMMIIOCKCNYN--CPHANEA 365
      |||
Db 919 -----QKCRCEHGAACDHVSGA 935

RESULT 11
T42215
zonadhesin - mouse

```

```

N:Alternate names: sperm-specific membrane protein
C:Species: Mus musculus (house mouse)
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C:Accession: T42215
R:Gao, Z.; Garbers, D.L.
J. Biol. Chem. 273, 3415-3421, 1998
A:Title: Species diversity in the structure of zonadhesin, a sperm-specific membrane
A:Reference number: Z22080; MUID:98123114; PMID:9452463
A:Accession: T42215
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-5376 <GAO>
A:Cross-references: EMBL:U07068; NID:g3327420; PID:g3327421; PIDN:AAC26680.1
C:Genetics:
A:Gene: Zan
A:Map position: 5
A:Function:
A>Description: functions in multiple cell adhesion processes
A:Note: found exclusively on the apical region of the sperm head
C:Keywords: cell adhesion

Query Match 7.28; Score 152; DB 2; Length 5376;
Best Local Similarity 21.08; Pred. No. 0.0082;
Matches 81; Conservative 39; Mismatches 138; Indels 128; Gaps 19;

OY 19 TRLALSTCPAACHCPLEAPKCAPGVGLVRDGCCKVCANQINDECKTOPCDHTKGLGEC 78
      |||
Db 4378 TKIALQ-CPAHSHTTYTLPSCITSCSNVNDRC-----ESTLQRTSTICGLC 4425
OY 79 NFGASSTALGICRAQ-----SEGRPEYNSRIYONGESF-QPNCQOCTCIDGANGCIP 132
      |||
Db 4426 HSGFVFEKDKCVPRTOGCKDSQ-----TLIPAKNMVTTTCGSCRCCTGSLVQCHD 4478
OY 133 L-CPQELSLPNLGCPNRLVKNYQC-----CE-----EMVCDSDSIKPM 173
      |||
Db 4479 FOCPSGAECDDIEDGNSNCVEITVQCPAHSYSKCLPPCQPSGSDPGHCEGTSPEAPST 4538
OY 174 DOGSLKELGEPDSEVELFRNNELIIVAGKRSILKRLPVFGMEPRILYNPLQOKCIYQT 233
      |||
Db 4539 CEECV-----CEPDY--LSNDKCVPS- 4559
OY 234 TSMQCKTKGCTGISTRVTNDNPECLRVKETRICEVRPC-----GQPVYS--SLKKG 283
      |||
Db 4560 ---SECCGKAHGV-----LIPESKTVMSGCKRKNCKCKGTGQCHDFSCPTG 4604
OY 284 KCKSKTKKSP-----PVRETYAGCLSVKKYRKYCGSCVD--GRC--CTPOLRT 330
      |||
Db 4605 SRCIDNNEGNSNCVTYALKCPAHSLYTNCL-----PCLPSCSDPGLCGTSPVPSI 4658
OY 331 VKARFRCEDETFPSKNYMIQ--SKC 355
      |||
Db 4659 CKEGCLCQSGVLAHKNCMLRIHCD 4684

RESULT 12
A39804
thrombospondin precursor - chicken
C:Species: Gallus gallus (chicken)
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: A39804
R:Lawler, J.; Dugnette, M.; Ferro, P.
J. Biol. Chem. 266, 8039-8043, 1991
A:Title: Cloning and sequencing of chicken thrombospondin.
A:Reference number: A39804; MUID:91217026; PMID:2022631
A:Accession: A39804
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1178 <LAW>
A:Cross-references: GB:M60853; NID:g212763; PIDN:AAA51437.1; PID:g212764
C:Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology;
F:325-383/Domain: von Willebrand factor type C repeat homology <VWC>
F:386-437/Domain: thrombospondin type 1 repeat homology <THRI>
F:442-498/Domain: thrombospondin type 1 repeat homology <THR2>

```

F;499-555/Domain: thrombospondin type 1 repeat homology <THR3>
F;658-697/Domain: EGF homology <EGF>

Query Match	7.1%;	Score 150.5;	DB 1;	Length 1178;
Best Local Similarity	23.0%;	Pred. No. 0.0025;		
Matches 65;	Conservative 33;	Mismatches 102;	Indels 83;	Gaps 13;

QY	100	CENSRIRIONGSEFCNCOHOCTCIDGAAGCIPLCFOELSLPLGGPNRLVAVTGOCC	159
Db	326	CMOGRVADBSRWIVDSTCKTCCTODSKIVC-----HQTCPVSCADSFE--GECCP	378
QY	160	EWVCDEDSIKDPMEDQDLGLKELGPDASEVELTRNNELLIAVGKGSILKLPIYFGEPR	219
Db	379	--VC---SHSDSER-----GWSPWMDMKCVTKCSGTQMGRRSDDVTRSACTG--PHI	426
QY	220	LYNPLQGKC---IYQT-----TSMQCKSTKCGTGISTRVNDNECRIV-----	261
Db	427	QTRMCSFFKCDHRITKODGGMHSWSPSSCVTCGVGNTRIRLNSPIDPMGSKNCVNG	486
QY	262	KETRICEVRPC-----	283
Db	487	RETERCEAPPCPVNQOMGPWSFWMSACTVTCGGIRERSRLCNSPEQYGAKPCVDJTKH	546
QY	284	KKCSKTTSPEEVRTFYTAGCLSYKKRYRKYGCSCVDGR-CTTP	325
Db	547	DMCKN-RDCP-----IDGLSNPCFPGAECNSYIPDGWSNCGP	582

RESULT 13

mucin-like peptide MLP 2677 - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 04-Mar-1993 #sequence:revision 18-Nov-1994 #text_change 10-Oct-1997
 C:Accession: A42112
 R:Xu, G.; Huan, L.J.; Khatri, I.A.; Wang, D.; Bennick, A.; Fahm, R.E.; Forstner, G.G.
 J. Biol. Chem. 267, 5401-5407, 1992
 A:Title: cDNA for the carboxyl-terminal region of a rat intestinal mucin-like peptide.
 A:Reference number: A42112; MUID:92184794; PMID:1377199

A:Status: preliminary
A:Molecule type: nucleic acid; protein
A:Residues: 1-837 <X01>
A:Experimental source: intestine
A:Note: sequence inconsistent with the nucleotide translation
A:Note: sequence extracted from NCBI backbone (NCBI:P:87420)
C:Superfamily: von Willebrand factor type C repeat homology
C:582-650/Domain: von Willebrand factor type C repeat homology <WVC>

Query Match	7.18;	Score 150;	DB 2;	length 837;
Best Local Similarity	19.28;	Pred. No. 0.0019;		
Matches	88;	Conservative	40;	Mismatches 138; Indels 192; Gaps 24;
QY	25	TCPA-----	ACHCPLEAPKCAPGVGLVADGGC-----	52
Db	419	TCFAHNOYQACGSPSEPTCOSSSPKKNSTLLVEGCFCEBGTTKAFAPGIDVCYKTCGCGPD		478
QY	53	-----CKVC-----	AKOLNEDCSKTOPCDH	72
Db	479	NVREGEHEFEPOKDCVCLGGSGIYCQPKKARGNLTTCEEDGTLYIVAADDDCCN		538
QY	73	TKGLECN-----	FGASSALKGICRAOSEGRP-----CEYNSRIYONGESFP	115
Db	539	TTSCCKDPKPKCAERPSCLLGFVEKSHVYGKCCPYSCVPKGYCVHNEAYIGSPVYS		598
QY	116	NCQHCQTCID-----GAVGC--IPLCQELSLPLGCPN--RLYKVTYGCCCEWYCD		164
Db	599	NKQDCQVCYCIDSMNSTQNLAVISCTHVP-C-----NISCSGFELVEPRECCCKKQQT		650
QY	165	EDSKDPMED-----ODGLLKEELG-----FDASEVELTTRNNELLAVAGSGIKRLPFGM		215
Db	651	HCILKRPEQGYILLRKEIGIKRNNDRCTFESCKKI-----NNQLIS-----SVSNITCPDF		701
QY	216	EPRLIXNPLDQKACIVQTTSW--SQCSKTCGTGISTRYTNDNPECLVKTETRICVEBPG		273

```

Db      702 DP-----SDCVGSIYMPNGGCKK-----IHNPR-----NTYRPS 733
      724 Q-FYISLKKGKKCKSTKKSPEPVRFTYAGCLSVKKYRKYC-GSC-----VD 318
      734 ALEVMKEI-----SYNGC--ARNISMNCAGSGCFAMYSAAODLD 773
QY      320 G--RCCTPOLRTVTKMRFCDEGETPFSKKVMVMIQSCK 355
      774 HGCSCCREKTSYRWVSLDCCPGSKLSHSTYTHIESCLC 811

```

RESULT 14

laminin B1k chain precursor - human
N:Alternate names: kalinin B1 chain
C:Species: Homo sapiens (man)
C:Date: 07-Oct-1994 #sequence_revision 07-Oct-1994 #text_change 05-Nov-1999
C:Accession: A53612
R:Gercke, D.R., Magman, D.W., Champliand, M.F., Burgess, R.E.
J. Biol. Chem. 269, 11073-11080, 1994
A:Title: The complete primary structure for a novel laminin chain, the laminin B1k chain
A:Reference number: A53612; MUID:94209274; PMID:7512558
A:Accession: A53612
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1170 <GER>
A:Cross-references: GB:I225541; NID:g510702; PIDN:AAA61834.1; PID:g510703
A:Note: authors translated the codon CGA for residue 124 as Gln, GAG for residue 439
C:Superfamily: laminin-type EGF-like homology
C:Keywords: glycoprotein
F:1-17/Domain: signal sequence #status predicted <SIG>
F:18-1170/Product: laminin B1k chain #status predicted <MAT>
F:250-312/Domain: laminin-type EGF-like homology <LEG1>
F:378-427/Domain: laminin-type EGF-like homology <LEG>
F:450-476/Domain: laminin-type EGF-like homology <EG11>
F:532-576/Domain: laminin-type EGF-like homology <LEG2>

Query Match	7.0%	Score 147.5	DB 2	Length 1170
Best Local Similarity	20.9%	Pred. No. 0.0041		
Matches 84	Conservative 31	Mismatches 101	Indels 185	Gaps 22

[illegible]

RESULT 15

109059

notch4 - mouse

C:Species: Mus musculus (house mouse)
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 08-Sep-2002
C:Accession: T09059
R:Rowen, L.; Mahairas, G.; Qin, S.; Ahearn, M.E.; Dankers, C.; Lasky, S.; Loretz, C.; Sc
submitted to the EMBL Data Library, October 1997
A:Description: Sequence of the mouse major histocompatibility locus class III region.
A:Reference number: 216543
A:Accession: T09059
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1964 <ROW>
A:Cross-references: EMBL:AF030001; NID:92564945; PID:92564947
C:Genetics:
A:Gene: notch4
A:Map position: 17
A:Introns: 22/1; 49/2; 148/1; 264/1; 305/1; 384/1; 436/1; 501/1; 539/1; 577/1; 618/1; 67
1679/3; 1729/1; 1761/3
C:Superfamily: notch protein; ankyrin repeat homology; EGF homology
C:Keywords: receptor; signal transduction
F:514-545/Domain: EGF homology <EGF>

Query Match 6.9%; Score 146; DB 2; Length 1964;
Best Local Similarity 23.4%; Pred. No. 0.0086;
Matches 94; Conservative 28; Mismatches 141; Indels 138; Gaps 24;

```
QY 28 AACGCPLEAPK--CAPGVGLVRDGGCCCKVCARQINEDCSKT-----OP-----CD 71
    |||  :  |||  :  |||  :  |||  :  |||  :  |||  :  |||  :  |||  :  |||  :
Db 523 AACHDILNGFQCLLPFTGAR-----CEKDME--CSSTPCANGRCRDQPAFYCE 573
    |||  :  |||  :  |||  :  |||  :  |||  :  |||  :  |||  :  |||  :  |||  :
QY 72 HTKGL-----CNFGASTALKG-----ICRAQSEGRPEY---NSRIYON 109
    |||  :  |||  :  |||  :  |||  :  |||  :  |||  :  |||  :  |||  :  |||  :
Db 574 CLPGEHPHCEKEVDECLDPCPVGASCLDLPFAFCLCRPGFTGQLCEVPLCTPNMCP 633
    |||  :  |||  :  |||  :  |||  :  |||  :  |||  :  |||  :  |||  :  |||  :
QY 110 GESFPNCOHQ-----CTCIDGAVGCIPL-----CPQELSLPNLGCPNRLVAVT 154
    |||  :  |||  :  |||  :  |||  :  |||  :  |||  :  |||  :  |||  :  |||  :
Db 634 GQ-----QCQGEHRAPCLPDGSFGCPAEDMCPCHHGCQSLCYCDEGWTGPCETEL 689
    |||  :  |||  :  |||  :  |||  :  |||  :  |||  :  |||  :  |||  :  |||  :
QY 155 GOCCEEWVCDDEDSIKDPMEDODGLGKELGFDASEVELTRNNELIYAVGKRSKRLPVFG 214
    |||  :  |||  :  |||  :  |||  :  |||  :  |||  :  |||  :  |||  :  |||  :
Db 690 GGCIST-PCAHGTCCHPOPS-----GYNCTCPAGYMGITCSEVTACHSGPC----- 735
    |||  :  |||  :  |||  :  |||  :  |||  :  |||  :  |||  :  |||  :  |||  :
QY 215 MEPRILYNPLQOKCTIVOTTSMSQCSKTGCTGISTVTNDNPECRVKEETRICE--VRPC 272
    |||  :  |||  :  |||  :  |||  :  |||  :  |||  :  |||  :  |||  :  |||  :
Db 736 -----LNGSGCSIREGYSCITLPSHTG-----RHQQTAVDHC 768
    |||  :  |||  :  |||  :  |||  :  |||  :  |||  :  |||  :  |||  :  |||  :
QY 273 GQPVYSLKKKKCKSKTKSPKSPKSPKSPKSPKSPKSPKSPKSPKSPKSPKSPKSPKSPK 322
    |||  :  |||  :  |||  :  |||  :  |||  :  |||  :  |||  :  |||  :  |||  :
Db 769 ---VSASCLNGTCVN-----KQTF--CLCATGFGGLHCEKTNFSCADSPCRNKAT 817
    |||  :  |||  :  |||  :  |||  :  |||  :  |||  :  |||  :  |||  :  |||  :
QY 323 C--TPQLTRYVKMRFCEDGETFSKNVMTQSCCKNTNCPH 361
    |||  :  |||  :  |||  :  |||  :  |||  :  |||  :  |||  :  |||  :  |||  :
Db 818 CQDTPRGARCL-----CSPGTGSSCQTLIDLQ-ARRCPH 852
    |||  :  |||  :  |||  :  |||  :  |||  :  |||  :  |||  :  |||  :  |||  :
```

Search completed: August 5, 2003, 14:07:29
Job time : 17.5408 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 5, 2003, 14:02:16 ; Search time 9.525 Seconds

(without alignments)
1881.066 Million cell updates/sec

Title: US-09-495-448A-4
Perfect score: 2116
Sequence: 1 MSRRIRALALVLTLLHTR.....ANEAPPYRLFNDHKFRD 381

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database: SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2106	99.5	381	1	CYR6_HUMAN
2	1929	91.2	379	1	CYR6_MOUSE
3	1929	91.2	379	1	CYR6_RAT
4	1695	80.1	375	1	CELO_CHICK
5	955	45.1	347	1	CTGF_MOUSE
6	952.5	45.0	346	1	CTGF_MOUSE
7	947	44.8	349	1	CTGF_MOUSE
8	933	44.1	349	1	CTGF_BOVIN
9	906	42.8	349	1	CTGF_MOUSE
10	851.5	40.2	351	1	NOV_MOUSE
11	848	40.1	353	1	NOV_MOUSE
12	827.5	39.1	357	1	NOV_MOUSE
13	825.5	39.0	343	1	NOV_MOUSE
14	822	38.8	354	1	NOV_MOUSE
15	818.5	38.7	351	1	NOV_MOUSE
16	523.5	24.7	250	1	CTGF_MOUSE
17	522.5	24.7	251	1	CTGF_MOUSE
18	490	23.2	250	1	CTGF_MOUSE
19	157	7.4	1700	1	BAR3_CHICK
20	153.5	7.3	2282	1	ZAN_RABIT
21	152	7.2	3110	1	LM33_HUMAN
22	152	7.2	5376	1	ZAN_MOUSE
23	150.5	7.1	1178	1	TSP2_MOUSE
24	150	7.1	837	1	MUC1_MOUSE
25	147.5	7.0	456	1	NRL1_MOUSE
26	146	6.9	1964	1	NTC4_MOUSE
27	144.5	6.8	1172	1	LM33_HUMAN
28	143	6.8	1173	1	TSP1_XENIA
29	141	6.7	480	1	HRA1_MOUSE
30	141	6.7	1233	1	MUC3_HUMAN
31	141	6.7	5179	1	MUC3_HUMAN
32	139	6.6	447	1	NRL1_MOUSE
33	139	6.6	1170	1	TSP1_BOVIN

ALIGNMENTS

RESULT 1	CYR6_HUMAN	STANDARD:	PRT:	381 AA.
AC	000622; 014934; 043775; 09BZL7;			
DT	15-JUL-1998 (Rel. 36, Created)			
DT	15-JUL-1998 (Rel. 36, Last sequence update)			
DT	15-SEP-2003 (Rel. 42, Last annotation update)			
DE	CYR61 protein precursor (Cysteine-rich, angiogenic inducer, 61)			
DE	(Insulin-like growth factor-binding protein 10) (IGFBP10 protein).			
GN	CYR61 OR IGFBP10 OR GIG1.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Albrecht C., von der Kammer H., Klaudiny J., Mayhaus M., Nitsch R.M.,			
RT	Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RA	MEDLINE=97280750; PubMed=9135077;			
RT	Jay P., Borge-Defranc J.L., Marsollier C., Mejean C., Tavlaux S.,			
RT	Berta P.,			
RT	"The human growth factor-inducible immediate early gene, CYR61, maps			
RT	to chromosome 1p."			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RA	MEDLINE=98197344; PubMed=9536281;			
RA	Martinerie C., Viegas-Pequignot E., Nguyen V.C., Perbal B.,			
RT	"Chromosomal mapping and expression of the human cyr61 gene in tumour			
RT	cells from the nervous system."			
RT	Mol. Pathol. 50:310-316(1997).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RA	RISSUE-Placenta;			
RA	Kolesnikova T.V., Lau L.F.,			
RT	Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.			
RN	[5]			
RP	SEQUENCE FROM N.A.			
RA	Bl A.B., Yu L.,			
RT	Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.			
RN	[6]			
RP	SEQUENCE FROM N.A.			
RA	Schuetz N., Lechner A., Groll C., Koehle J., Jakob F.,			
RT	"Regulation of hCYR61 by vitamin D, serum and cytokines in fetal human			
RT	osteoblasts."			
RT	Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.			
RN	[7]			
RP	SEQUENCE FROM N.A.			
RA	Ieng E., Tai G., Estable M., Liu J., Chow C., Sadowski I.,			
RT	"Organization and expression of the CYR61 gene in normal human			
RT	fibroblasts."			
RT	Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.			
RN	[8]			
RP	SEQUENCE FROM N.A.			

34	139	6.6	1587	1	LMG3_HUMAN	O9606 homo sapien
35	139	6.6	2437	1	NTC1_BRARE	P46530 brachydanio
36	139	6.6	2482	1	WVF_PIG	Q28833 sus scrofa
37	138.5	6.5	2703	1	NOTC2_MOUSE	P07207 drosophila
38	137.5	6.5	1504	1	SLIT1_MOUSE	P24014 drosophila
39	137	6.5	453	1	HRA3_HUMAN	P83110 homo sapien
40	137	6.5	870	1	SRG2_HUMAN	O96966 homo sapien
41	137	6.5	2531	1	NTC1_MOUSE	O07008 rattus norv
42	136.5	6.5	833	1	SRG2_MOUSE	P59222 mus musculu
43	135	6.4	1168	1	LM33_MOUSE	O61087 mus musculu
44	135	6.4	1170	1	TSP1_MOUSE	P35441 mus musculu
45	135	6.4	2813	1	WVF_CANPA	Q28295 canis fami1

RC TISSUE-Lung, Placenta, and Skin;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshlyuk S., Carroll P., Prange C.,
 RA Raba S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunatirne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hallyk S.W.,
 RA Villalón D.K., Muzny D.C., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butlerfield T.S.N., Krzywinski M.T., Skalska U., Smallus D.E.,
 RA Schnerich A., Schein J.E., Jones S.J.M., Maira M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -1- FUNCTION: MAY ACT AS ONE OF THE MANY GROWTH FACTOR-BINDING
 CC PROTEINS. PROMOTES PROLIFERATION, MIGRATION AND ADHESION (BY
 CC SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: Contains 1 IGFBP domain.
 CC -1- SIMILARITY: Contains 1 WFEC domain.
 CC -1- SIMILARITY: Contains 1 TSP type-1 domain.
 CC -1- SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed, usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; Y12084; CAI72802.1; -;
 DR EMBL; U62015; AAB58319.1; -;
 DR EMBL; Y11307; CAI72167.1; -;
 DR EMBL; AF003594; AAB61240.1; -;
 DR EMBL; AF031385; AAB84227.1; -;
 DR EMBL; Z98053; CAB10848.1; -;
 DR EMBL; AF307860; AAG59863.1; -;
 DR EMBL; BC001271; AAH01271.1; -;
 DR EMBL; BC009199; AAH09199.1; -;
 DR EMBL; BC016952; AAH16952.1; -;
 DR GeneW; HGNC:2634; CYR61.
 DR MIM; 602369; -;
 DR GO; GO:0008283; P:cell proliferation; TAS.
 DR GO; GO:0007345; P:embryogenesis and morphogenesis; TAS.
 DR InterPro; IPR006208; Cys_knot.
 DR InterPro; IPR006207; Cys_knot_C.
 DR InterPro; IPR000867; Insu_gro_fac_pr.
 DR InterPro; IPR000884; TSP1.
 DR Pfam; PF00007; Cys_knot_C.
 DR Pfam; PF000219; IGFBP.1.
 DR Pfam; PF00090; TSP1.1.
 DR Pfam; PF00093; vwc.1.
 DR SMART; SM00041; CT; 1.
 DR SMART; SM00121; IB; 1.
 DR SMART; SM00209; TSP1.1.
 DR SMART; SM00214; VWC.1.
 DR PROSITE; PS01185; CTCK_1; 1.
 DR PROSITE; PS01225; CTCK_2; 1.
 DR PROSITE; PS00222; IGF_BINDING; 1.
 DR PROSITE; PS00992; TSP1.1.
 DR PROSITE; PS01208; WFEC_1; 1.
 DR PROSITE; PS0184; WFEC_2; 1.
 KW Growth factor binding; Signal.

FT SIGNAL 1 24 POTENTIAL.
 FT CHAIN 25 381 CYR61 PROTEIN.
 FT DOMAIN 26 97 IGFBP.
 FT DOMAIN 98 164 WFEC.
 FT DOMAIN 228 273 TSP TYPE-1.
 FT DOMAIN 286 360 CTCK.
 FT DISULFID 286 323 BY SIMILARITY.
 FT DISULFID 303 337 BY SIMILARITY.
 FT DISULFID 314 353 BY SIMILARITY.
 FT DISULFID 317 355 BY SIMILARITY.
 FT DISULFID 322 359- BY SIMILARITY.
 FT CONFLICT 165 165 E -> Q (IN REF. 3).
 FT CONFLICT 210 210 L -> I (IN REF. 5).
 FT CONFLICT 220 220 F -> R (IN REF. 5).
 FT CONFLICT 369 369 F -> L (IN REF. 7).
 SO SEQUENCE 381 AA; 42026 MM; FC0BD39C078CA0B1 CRC64;
 Query Match 99.5%; Score 2106; DB 1; Length 381;
 Best Local Similarity 99.5%; Pred. No. 9.8e-146;
 Matches 379; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 DB 1 MSSRIARALAVTLHLTRIALSTCPACHCPLPAPKADGVLVRDGGCGCKVCAKOL 60
 1 MSSRIARALAVTLHLTRIALSTCPACHCPLPAPKADGVLVRDGGCGCKVCAKOL 60
 61 NEDCSKTOPCDHTKGLKCNFGASSTALKGICRAOSEGRPCERYNSRTYONGSEFOPNCOHQ 120
 61 NEDCSKTOPCDHTKGLKCNFGASSTALKGICRAOSEGRPCERYNSRTYONGSEFOPNCOHQ 120
 121 CTCIDGAVGCIPLCPQOEIPLNLGCPNRLVKVGGCCCEEWVCDSDSKIDMEODDGLG 180
 121 CTCIDGAVGCIPLCPQOEIPLNLGCPNRLVKVGGCCCEEWVCDSDSKIDMEODDGLG 180
 181 KELGDAEVEELTRNNELIANGKSKRLPVGNEPRILYNPLGOKACIYOTTSWQCS 240
 181 KELGDAEVEELTRNNELIANGKSKRLPVGNEPRILYNPLGOKACIYOTTSWQCS 240
 181 KELGDAEVEELTRNNELIANGKSKRLPVGNEPRILYNPLGOKACIYOTTSWQCS 240
 241 KTCGIGSTRYTNNDNPEKRLKERRICVRCGQPVYSSLLKGGKSKTSKSPERYRTY 300
 241 KTCGIGSTRYTNNDNPEKRLKERRICVRCGQPVYSSLLKGGKSKTSKSPERYRTY 300
 241 KTCGIGSTRYTNNDNPEKRLKERRICVRCGQPVYSSLLKGGKSKTSKSPERYRTY 300
 301 AGCLSVKKRKYKGYGSCVDGRCQPOLRRTYKMRPCDGEFENVMYIOSCKNVCNP 360
 301 AGCLSVKKRKYKGYGSCVDGRCQPOLRRTYKMRPCDGEFENVMYIOSCKNVCNP 360
 301 AGCLSVKKRKYKGYGSCVDGRCQPOLRRTYKMRPCDGEFENVMYIOSCKNVCNP 360
 361 HANEAPFPYRLFNDIHKFRD 381
 361 HANEAPFPYRLFNDIHKFRD 381
 DB 361 HANEAPFPYRLFNDIHKFRD 381
 DB 361 HANEAPFPYRLFNDIHKFRD 381
 RESULT 2
 CYR6_MOUSE STANDARD; PRT; 379 AA.
 ID CYR6_MOUSE STANDARD; PRT; 379 AA.
 AC P18406;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE CYR61 protein precursor (Cysteine-rich, angiogenic inducer, 61)
 DE (Insulin-like growth factor-binding protein 10) (3CH61).
 GN CYR61 OR IGFBP10.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP STRAIN-BALB/C; TISSUE=Fibroblast;
 RC MEDLINE=90287146; PubMed=2355916;
 RX O'Brien T.P., Yang G.P., Sanders L., Lau L.F.;
 RT "Expression of cyr61, a growth factor-inducible immediate-early
 RT gene.";
 RL Mol. Cell. Biol. 10:3569-3577(1990).
 RN [2]

RP SEQUENCE FROM N.A.
 RC STRAIN-AJ; TISSUE=Embryonic fibroblast;
 RX MEDLINE=91288203; PubMed=2062642;
 RA Latickie B.V., O'Brien T.P., Lau L.F.;
 RT "Promoter function and structure of the growth factor-inducible
 RT Immediate early gene cyr61."
 RL Nucleic Acids Res. 19:3261-3267(1991).
 CC -1- FUNCTION: MAY ACT AS ONE OF THE MANY GROWTH FACTOR-BINDING
 CC PROTEINS; PROMOTES PROLIFERATION, MIGRATION AND ADHESION.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: LOW IN KIDNEY, ADRENAL GLAND, TESTES, BRAIN,
 CC AND OVARY, MODERATE IN HEART, UTERUS, AND SKELETAL MUSCLE, HIGHEST
 CC IN LUNG.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED FROM G(0)/G(1) THROUGH MID-G(1) IN
 CC NORMAL CELLS, AND AT A CONSTANT LEVEL IN RAPIDLY GROWING CELLS.
 CC -1- INDUCTION: By growth factors.
 CC -1- SIMILARITY: Contains 1 IGFBP domain.
 CC -1- SIMILARITY: Contains 1 WFPC domain.
 CC -1- SIMILARITY: Contains 1 TSP type-1 domain.
 CC -1- SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M32490; AAA37512.1; -;
 DR EMBL; X56790; CAA40109.1; -;
 DR PIR; A35669; A35669;
 DR MGI; MGI:88613; Cyr61.
 DR GO; GO:0001569; Patterning of blood vessels; IMP.
 DR InterPro; IPR006208; Cys_knot.
 DR InterPro; IPR006207; Ins1_knot_C.
 DR InterPro; IPR000867; Ins1_gro_fac_pr.
 DR InterPro; IPR000884; TSP1.
 DR InterPro; IPR001007; WFC_C.
 DR Pfam; PF00007; Cys_knot; 1.
 DR Pfam; PF00219; IGFBP; 1.
 DR Pfam; PF00090; TSP_1; 1.
 DR Pfam; PF00093; WFC; 1.
 DR SMART; SM00041; CT; 1.
 DR SMART; SM00121; IB; 1.
 DR SMART; SM00209; TSP1; 1.
 DR SMART; SM00214; WFC; 1.
 DR PROSITE; PS01185; CTCK_1; 1.
 DR PROSITE; PS01225; CTCK_2; 1.
 DR PROSITE; PS00223; IGF_BINDING; 1.
 DR PROSITE; PS50093; TSP1; 1.
 DR PROSITE; PS01208; WFC_1; 1.
 DR PROSITE; PS50184; WFC_2; 1.
 KW Growth factor binding; Signal.
 FT SIGNAL 1 24 POTENTIAL.
 FT CHAIN 25 379 CYR61 PROTEIN.
 FT DOMAIN 98 164 WFPC.
 FT DOMAIN 226 271 TSP TYPE-1.
 FT DOMAIN 284 358 CTCK.
 FT DISULFID 284 321 BY SIMILARITY.
 FT DISULFID 301 335 BY SIMILARITY.
 FT DISULFID 312 351 BY SIMILARITY.
 FT DISULFID 315 353 BY SIMILARITY.
 FT DISULFID 320 357 BY SIMILARITY.
 SQ SEQUENCE 379 AA; 41709 MM; FA6B501AB56A8EE9 CRC64;

Query Match 91.2%; Score 1929; DB 1; Length 379;
 Best Local Similarity 90.9%; Pred. No. 6.5e-133;
 Matches 348; Conservative 10; Mismatches 19; Indels 6; Gaps 2;

OY 1 MSRIARALAVVTLHLTRALSTCPACHCPLEAPKAPGVGLVDDGGCCGVCAKOL 60
 DB 1 MSSSTFRLAVAVTLHLTRALSTCPACHCPLEAPKAPGVGLVDDGGCCGVCAKOL 60
 OY 61 NEDSKTOPCDHTGKLECNFGASSTALKGICRAOSEGRPCYNSRIYONGESFQPNCOHQ 120
 DB 61 NEDSKTOPCDHTGKLECNFGASSTALKGICRAOSEGRPCYNSRIYONGESFQPNCOHQ 120
 OY 121 CTCIDGAVGCIPLCPQELSLPNLCGPNRLVKTGQCCSEWVCDESDIKDPMEDQDGLG 180
 DB 121 CTCIDGAVGCIPLCPQELSLPNLCGPNRLVKTGQCCSEWVCDESDIKDPMEDQDGLG 180
 OY 121 CTCIDGAVGCIPLCPQELSLPNLCGPNRLVKTGQCCSEWVCDESDIKDPMEDQDGLG 180
 DB 121 CTCIDGAVGCIPLCPQELSLPNLCGPNRLVKTGQCCSEWVCDESDIKDPMEDQDGLG 180
 OY 181 KELGFDSSEVLTNNELIANGKRSKLRLPFVEMERILYNPL--QGOKIVOTTSMSQ 238
 DB 179 --LGLDSEVELTNNELIALGKSSSLRLPFVETEPRLVLPNLAHQKCIYVOTTSMSQ 236
 OY 239 CSKTCGAGISRRVNDPECLVETRICERPCGOPYSSLKKGKSKTKSPPEYRF 298
 DB 237 CSKTCGAGISRRVNDPECLVETRICERPCGOPYSSLKKGKSKTKSPPEYRF 296
 OY 299 TYAGCLSVKKYRPYCGSCVDGRCCTPOLRTVYMRRCDEGTFSKNVMIOGCKNYN 358
 DB 297 TYAGCSSVKKYRPYCGSCVDGRCCTPOLRTVYMRRCDEGTFSKNVMIOGCKNYN 356
 OY 359 CPANEAPFPYRLPFNDIHKFRD 381
 DB 357 CPHPNEASFRLYSLFNDIHKFRD 379
 RESULT 3
 CYR6_RAT STANDARD; PRT; 379 AA.
 ID CYR6_RAT
 AC Q9ES72;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE CYR61 protein precursor (Cysteine-rich, angiogenic inducer, 61)
 DE (Insulin-like growth factor-binding protein 10).
 GN CYR61 OR IGFBP10.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lung;
 RX MEDLINE=20435857; PubMed=10852911;
 RA Albrecht C., von Der Kammer H., Mayhaus M., Klautz J., Schweizer M.,
 RT Nitsch R.M.;
 RT "Muscarinic acetylcholine receptors induce the expression of the
 RT Immediate early growth regulatory gene Cyr61."
 RL J. Biol. Chem. 275:28929-28936(2000).
 CC -1- FUNCTION: MAY ACT AS ONE OF THE MANY GROWTH FACTOR-BINDING
 CC PROTEINS; PROMOTES PROLIFERATION, MIGRATION AND ADHESION (By
 CC similarity).
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: Contains 1 IGFBP domain.
 CC -1- SIMILARITY: Contains 1 WFPC domain.
 CC -1- SIMILARITY: Contains 1 TSP type-1 domain.
 CC -1- SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF218568; AAG14964.1; -;
 DR InterPro; IPR006208; Cys_knot.
 DR InterPro; IPR006207; Cys_knot_C.
 DR InterPro; IPR000867; Ins1_gro_fac_pr.
 DR InterPro; IPR000884; TSP1.
 DR InterPro; IPR001007; WFC_C.
 DR Pfam; PF00007; Cys_knot; 1.

DR Pfam: PF00219; IGFBP, 1.
 DR Pfam: PF00090; tsp_1; 1.
 DR Pfam: PF00093; vwc; 1.
 DR SMART: SM00041; CT; 1.
 DR SMART: SM00121; IB; 1.
 DR SMART: SM00209; TSP1; 1.
 DR PROSITE: PS01185; CTCK; 1.
 DR PROSITE: PS01225; CTCK_2; 1.
 DR PROSITE: PS00222; IGF BINDING; 1.
 DR PROSITE: PS50092; TSP1; 1.
 DR PROSITE: PS01208; vwc; 1; 1.
 DR PROSITE: PS50184; vwc_2; 1.
 DR Growth factor binding; Signal.
 FT SIGNAL 1 24
 FT CHAIN 25 379
 FT DOMAIN 26 97
 FT DOMAIN 98 164
 FT DOMAIN 226 271
 FT DOMAIN 284 358
 FT DISULFID 284 321
 FT DISULFID 301 335
 FT DISULFID 312 351
 FT DISULFID 315 353
 FT DISULFID 320 357
 SQ SEQUENCE 379 AA; 41687 MW; 62BF0BBA4C5AFDE9 CRC64;

Query Match Best Local Similarity 91.2%; Score 1929; DB 1; Length 379;
 Matches 348; Conservative 10; Mismatches 19; Indels 6; Gaps 2;

OY 1 MSSRIARALAVVTLHLTRLALSTCPACHCPLPAPKAPGVGLVRDGGCCGCKVCAROL 60
 DB 1 MSSSTKTLAVAVTLHLTRLALSTCPACHCPLPAPKAPGVGLVRDGGCCGCKVCAROL 60
 OY 61 NEDCSKTQPCDHTKGLGECNFGASSALKGICRAOSBGRCEVNSRYONGESFOPKCAQ 120
 DB 61 NEDCSKTQPCDHTKGLGECNFGASSALKGICRAOSBGRCEVNSRYONGESFOPKCAQ 120
 OY 121 CTCIDGAVGCIPLCPQELSLPMLGCPNPRVLVVTGQCCCEWVCDDESDINDPMDODGLG 180
 DB 121 CTCIDGAVGCIPLCPQELSLPMLGCPNPRVLVVTGQCCCEWVCDDESDINDPMDODGLG 180
 OY 121 CTCIDGAVGCIPLCPQELSLPMLGCPNPRVLVVTGQCCCEWVCDDESDINDPMDODGLG 178
 DB 121 CTCIDGAVGCIPLCPQELSLPMLGCPNPRVLVVTGQCCCEWVCDDESDINDPMDODGLG 178
 OY 181 KLGFDASEVELTRNNELIAGKGSRLKRLPVFGMEPRILYNPL--QGCKIVQTTSMQ 238
 DB 179 --LGFDASEVELTRNNELIAGKGSRLKRLPVFGMEPRILYNPL--QGCKIVQTTSMQ 236
 OY 239 CSKTCGTGISTVTNDNPECRVLVETRICVPRCCQPVYSSLAKKKCKSKTKSPVPVF 298
 DB 237 CSKTCGTGISTVTNDNPECRVLVETRICVPRCCQPVYSSLAKKKCKSKTKSPVPVF 296
 OY 239 TYAGCLVYKRYKRYCGSCVDGRCCPOLTRTKKRFRCDEDEFTSKNMVIOQCKCNV 358
 DB 237 TYAGCSSYKRYKRYKRYCGSCVDGRCCPOLTRTKKRFRCDEDEFTSKNMVIOQCKCNV 356
 OY 359 CPHANEAAFPYRFLNDIHKFRD 381
 DB 357 CPHNEASFRLYSLFNDIHKFRD 379

RESULT 4
 CE10_CHICK STANDARD; PRT; 375 AA.
 AC P19336;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE CEF-10 protein precursor.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;

[1]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=89145206; PubMed=2537491;
 RX Simmons D.L., Levy D.B., Yannoni Y., Erikson R.L.,
 RA Identification of a phorbol ester-repressible v-src-inducible gene.";
 RT Proc. Natl. Acad. Sci. U.S.A. 86:1178-1182(1989).
 CC -1- FUNCTION: PROBABLE SECRETED REGULATORY PROTEIN.
 CC -1- SUBCELLULAR LOCATION: Secreted (Probable).
 CC -1- INDUCTION: BY V-SRC.
 CC -1- SIMILARITY: Contains 1 IGFBP domain.
 CC -1- SIMILARITY: Contains 1 vwc domain.
 CC -1- SIMILARITY: Contains 1 TSP type-1 domain.
 CC -1- SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

DR EMBL: J04486; AAA8661.1; -
 DR PIR: A41428; A41428.
 DR InterPro: IPR006208; Cys_knot.
 DR InterPro: IPR006207; Cys_knot_C.
 DR InterPro: IPR000867; Ins1_gro_fac_pr.
 DR InterPro: IPR001007; vwc_C.
 DR InterPro: IPR00007; Cys_knot; 1.
 DR Pfam: PF00219; IGFBP; 1.
 DR Pfam: PF00090; tsp_1; 1.
 DR Pfam: PF00093; vwc; 1.
 DR SMART: SM00041; CT; 1.
 DR SMART: SM00121; IB; 1.
 DR SMART: SM00209; TSP1; 1.
 DR SMART: SM00214; vwc; 1.
 DR PROSITE: PS01185; CTCK; 1; 1.
 DR PROSITE: PS01225; CTCK_2; 1.
 DR PROSITE: PS00222; IGF BINDING; 1.
 DR PROSITE: PS50092; TSP1; 1.
 DR PROSITE: PS01208; vwc; 1; 1.
 DR PROSITE: PS50184; vwc_2; 1.
 DR Growth factor binding; Signal.
 KW SIGNAL 1 22
 FT CHAIN 23 375
 FT DOMAIN 23 375
 FT DOMAIN 98 164
 FT DOMAIN 223 268
 FT DOMAIN 281 318
 FT DISULFID 281 318
 FT DISULFID 298 332
 FT DISULFID 309 348
 FT DISULFID 312 350
 FT DISULFID 317 354
 SQ SEQUENCE 375 AA; 40651 MW; 95F28533BE3505AE CRC64;

Query Match Best Local Similarity 81.1%; Score 1695; DB 1; Length 375;
 Matches 311; Conservative 80.0%; Pred. No. 5 8e-116; Mismatches 41; Indels 12; Gaps 7;

OY 1 MSSRIARALAVVTLHLTRLALSTCPACHCPLPAPKAPGVGLVRDGGCCGCKVCAROL 59
 DB 1 MSSAGARP-ALAAALICLARIALGSPCPAVCCCPAAAPQACGVLVDPGSCCCRCVCAKQ 59
 OY 60 LNEDCSKQPCDHTKGLGECNFGASSALKGICRAOSBGRCEVNSRYONGESFOPKCAQ 119
 DB 60 LNEDCSKQPCDHTKGLGECNFGASSALKGICRAOSBGRCEVNSRYONGESFOPKCAQ 119
 OY 120 OCTCIDGAVGCIPLCPQELSLPMLGCPNPRVLVVTGQCCCEWVCDDESDINDPMDODGLG 179
 DB 120 OCTCIDGAVGCIPLCPQELSLPMLGCPNPRVLVVTGQCCCEWVCDDESDINDPMDODGLG 177
 OY 180 GKLGFDASEVELTRNNELIAGKGSRLKRLPVFGMEP--RLYNPLQGGCKIVQTTSMQ 237

```

Db 178 SKERGLDASEGELTRNNELIAIYVG-GLKMLPVFGSEPSORAFENP-----KCIYQTSMS 232
QY 238 OCSKTGCTGISTRTVNDNPECRVLKETRICEVRRCGPVYSLSLKKGKSKTKKSPSPVR 297
Db 233 QCSKTGCTGISTRTVNDNPECRVLKETRICEVRRCGPVYSLSLKKGKSKTKKSPSPVR 292
QY 298 FTYAGCLSVKRRPKYKCGSCVDGRCTPOLTRVVKMRFRCEDETFESKNVMIOSCKNY 357
Db 293 FTYAGCSSLVKRRPKYKCGSCVDGRCTPOLTRVVKMRFRCEDETFESKNVMIOSCKNY 352
QY 358 NCPHANEAPFPYRLFNDFIKFRD 381
Db 353 NCPHANEAPFPYRLFNDFIKFRD 375

RESULT 5
CTGF_MOUSE STANDARD; PRT; 347 AA.
ID OQRIE9; 09WVS1;
AC 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE Connective tissue growth factor precursor.
GN CTGF.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP MEDLINE=20145935; PubMed=10679821;
RX Xu J., Smock S.L., Safadi F.F., Rosenzweig A.B., Odgren P.R.,
RA Marks S.C. Jr., Owen T.A., Popoff S.N.:
RT "Cloning the full-length cDNA for rat connective tissue growth factor:
RT implications for skeletal development."
RT J. Cell. Biochem. 77:103-115(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Tezuka K., Tamatani T.:
RT "Rattus norvegicus connective tissue growth factor."
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: MAJOR CONNECTIVE TISSUE MITOATPACTANT SECRETED BY
CC VASCULAR ENDOTHELIAL CELLS. PROMOTES PROLIFERATION AND
CC DIFFERENTIATION OF CHONDROCYTES (BY SIMILARITY). MEDIATES CELL
CC ADHESION AND ENHANCES FIBROBLAST GROWTH FACTOR-INDUCED DNA
CC SYNTHESIS (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: Contains 1 IGFBP domain.
CC -1- SIMILARITY: Contains 1 WFCC domain.
CC -1- SIMILARITY: Contains 1 TSP type-1 domain.
CC -1- SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AF120275; AAD39132.1; -.
CC DR EMBL: AB023068; BAA82125.1; -.
CC DR InterPro: IPR006208; Cys_knot.
CC DR InterPro: IPR006207; Cys_knot_C.
CC DR InterPro: IPR000867; Ins1_gro_fac_pr.
CC DR InterPro: IPR000884; TSP1.
CC DR InterPro: IPR001007; WMF_C.
CC DR Pfam: PF00007; Cys_knot; 1.
CC DR Pfam: PF000219; IGFBP; 1.
CC DR Pfam: PF00030; TSP_1; 1.
CC DR Pfam: PF00093; vwc; 1.
CC SMART: SM00041; CT; 1.

```

```

DR SMART: SM00121; IB; 1.
DR SMART: SM00209; TSP1; 1.
DR SMART: SM00214; vwc; 1.
DR PROSITE: PS01185; CTCK_1; 1.
DR PROSITE: PS01285; CTCK_2; 1.
DR PROSITE: PS00222; IGF_BINDING; 1.
DR PROSITE: PS50092; TSP1; 1.
DR PROSITE: PS01208; WMF_C; 1.
DR PROSITE: PS50184; WMF_C_2; 1.
DR Cell adhesion; DNA synthesis; Extracellular matrix; Signal.
KM SIGNAL 1 24
FT CHAIN 1 25 347
FT DOMAIN 31 98
FT DOMAIN 99 165
FT DOMAIN 196 241
FT DOMAIN 234 328
FT DISULFID 254 291
FT DISULFID 271 305
FT DISULFID 282 321
FT DISULFID 285 323
FT DISULFID 290 327
FT CONFLICT 35 35
FT CONFLICT 94 94
SQ SEQUENCE 347 AA; 37756 MW; CFBELAI976687B16 CRC64;

Query Match 45.18; Score 955; DB 1; Length 347;
Best Local Similarity 46.28; Pred. No. 2.2e-62;
Matches 176; Conservative 58; Mismatches 107; Indels 40; Gaps 7;

QY 1 MSSRIARALALVYTLHLTRIAL-STCPACHQPLE-APKCAPGVGLVRDCCGCKYCAK 58
Db 1 MASVAGPVSLALVLLCTRATGQDCSAGQCAAEAPRCAPGVSLVLDGCGCCRCYCAK 60
QY 59 QLNEDCSKTPCCHTKGLECNFASSTALKGICRAQSEGRPCENRHYNGSFQPNQC 118
Db 61 QLELCETERPDCPHKGLFCDFGSPANRKGICVCTAK-DGAPCVFGSVYSGESFQSSCK 119
QY 119 HOCCTCIDGAVGCIPLCPQELPNIPLGCPNRLVKNVGGCCCEWVCDEDSIKDMEDQDL 178
Db 120 YQCTCIDGAVGCVPLCSMDVRLPSDCPPRRVRLPKCCCEWVCDEP----- 167
QY 179 LGEELGFDASEVELTRNNELIAVKGKSLRL-PVQMEPRILYNPLQGGCKIVQTSMS 237
Db 168-----KDTIVYGPAALARLDTREPDPTM-----RANCLVQTSMS 205
QY 238 QCSKTGCTGISTRTVNDNPECRVLKETRICEVRRCGPVYSLSLKKGKSKTKKSPSPVR 297
Db 206 ACSKTGCMGISTRTVNDNPECRVLKETRICEVRRCGPVYSLSLKKGKSKTKKSPSPVR 265
QY 298 FTYAGCLSVKRRPKYKCGSCVDGRCTPOLTRVVKMRFRCEDETFESKNVMIOSCKNY 357
Db 266 FELSGCTSVYTRAKFCGCTDGRCCPHPTTLVPEFKCPDDEIMKKNMFKTCAHY 325
QY 358 NCPHANEAPFPY--RLFNDFI 376
Db 326 NCPGDMNDIFESLYRKMYGDM 346

RESULT 6
CTGF_MOUSE STANDARD; PRT; 348 AA.
ID AC P29268; 1992U0;
AC 01-DEC-1992 (Rel. 24, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Connective tissue growth factor precursor (FISP-12 protein)
DE (Hypertrophic chondrocyte-specific protein 24).
GN CTGF OR FISP12 OR FISP-12 OR HCS24.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]

```

RP SEQUENCE FROM N.A.
RX MEDLINE-91363290; PubMed-1886698;
RA Rybeck R.-P., Macdonald-Bravo H., Mattei M.-G., Bravo R.;
RT "Structure, mapping, and expression of flsp-12, a growth factor-
RT inducible gene encoding a secreted cysteine-rich protein.";
RT Cell Growth Differ. 2:225-233(1991).
[2]
RP SEQUENCE FROM N.A.
RX MEDLINE-91229699; PubMed-2029337;
RA Brunner A., Chini J., Neubauer M.G., Purchio A.F.;
RT "Identification of a gene family regulated by transforming growth
RT factor-beta.";
RT DNA Cell Biol. 10:293-300(1991).
[3]
RP SEQUENCE FROM N.A.
RX MEDLINE-22388257; PubMed-12477932;
RA Klausner R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shemmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Ditchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carinci P., Prange C.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Botterfield Y.S.N., Krzywinski M.I., Skalski U., Smallus D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[4]
RP FUNCTION, AND SUBCELLULAR LOCATION.
RX MEDLINE-97327410; PubMed-9184077;
RA Kireeva M.L., Latifkic B.V., Kolesnikova T.V., Chen C.C., Yang G.P.,
RA Ahler A.S., Lau L.F.;
RT "Cyt61 and flsp12 are both ECM-associated signalling molecules:
RT activities, metabolism, and localization during development.";
RT Exp. Cell Res. 233:63-77(1997).
[5]
RP FUNCTION.
RX MEDLINE-99182484; PubMed-10082563;
RA Babic A.M., Chen C.C., Lau L.F.;
RT "Flsp12/mouse connective tissue growth factor mediates endothelial
RT cell adhesion and migration through integrin alphavbeta3, promotes
RT endothelial cell survival, and induces angiogenesis in vivo.";
RT Mol. Cell. Biol. 19:2958-2966(1999).
CC -1- FUNCTION: MAJOR CONNECTIVE TISSUE MITOATTRACTANT SECRETED BY
CC VASCULAR ENDOTHELIAL CELLS. PROMOTES PROLIFERATION AND
CC DIFFERENTIATION OF CHONDROCYTES (BY SIMILARITY). MEDIATES CELL
CC ADHESION AND ENHANCES FIBROBLAST GROWTH FACTOR-INDUCED DNA
CC SYNTHESIS (BY SIMILARITY).
CC -1- SUBUNIT: Monomer (By similarity).
CC -1- SUBCELLULAR LOCATION: FOUND IN THE EXTRACELLULAR MATRIX AND AS A
CC SOLUBLE FORM.
CC -1- TISSUE SPECIFICITY: TESTIS, SPLEEN, KIDNEY, LUNG, HEART, AND BRAIN
CC (LOWEST LEVEL IN TESTIS AND HIGHEST IN LUNG).
CC -1- INDUCTION: By growth factors.
CC -1- SIMILARITY: Contains 1 IGFBP domain.
CC -1- SIMILARITY: Contains 1 VWF domain.
CC -1- SIMILARITY: Contains 1 TSP type-1 domain.
CC -1- SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation on
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

CC entries requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: M70641; AAA37627.1; -
DR EMBL: M70642; AAA37628.1; -
DR EMBL: M80263; AAA73135.1; -
DR EMBL: BC006783; AAH06783.1; -
DR PIR: A40578; A40578.
DR MGD: MGI:95537; Ctgf.
DR GO: GO:0005578; C:extracellular matrix; IDA.
DR GO: GO:0008201; F:heparin binding activity; IDA.
DR GO: GO:0005178; F:integrin binding activity; IDA.
DR GO: GO:0001525; P:angiogenesis; IDA.
DR GO: GO:0016477; P:cell migration; IDA.
DR GO: GO:0001610; P:cell-cell adhesion; IDA.
DR GO: GO:0008543; P:FGF receptor signaling pathway; IDA.
DR GO: GO:0007229; P:integrin-mediated signaling pathway; IDA.
DR GO: GO:0001503; P:osification; IMP.
DR InterPro: IPR006208; Cys_knot.
DR InterPro: IPR006207; Cys_knot_C.
DR InterPro: IPR000867; Ins1_gro_fac-pr.
DR InterPro: IPR001007; VWF_C.
DR Pfam: PF00007; Cys_knot.1.
DR Pfam: PF00219; IGFBP.1.
DR Pfam: PF00090; TSP_1.1.
DR Pfam: PF00093; VWC.1.
DR SMART: SM00041; CT.1.
DR SMART: SM00121; IB.1.
DR SMART: SM00209; TSP1.1.
DR SMART: SM00214; VWC.1.
DR PROSITE: PS01185; CTCK.1.
DR PROSITE: PS01225; CTCK_2.1.
DR PROSITE: PS00222; IGF_BINDING.1.
DR PROSITE: PS50092; TSP1.1.
DR PROSITE: PS01208; VWF_C.1.
DR PROSITE: PS01208; VWF_C_2.1.
KW Cell adhesion; DNA synthesis; Extracellular matrix; Signal.
FT CHAIN 1 25 POTENTIAL.
FT DOMAIN 26 348 CONNECTIVE TISSUE GROWTH FACTOR.
FT DOMAIN 32 99 IGFBP.
FT DOMAIN 100 166 VWF.
FT DOMAIN 197 242 TSP TYPE-1.
FT DOMAIN 255 329 CTCK.
FT DISULFID 255 292 BY SIMILARITY.
FT DISULFID 272 306 BY SIMILARITY.
FT DISULFID 283 322 BY SIMILARITY.
FT DISULFID 286 324 BY SIMILARITY.
FT DISULFID 291 328 BY SIMILARITY.
FT CONFLICT 161 161 E -> K (IN REF. 1).
SQ SEQUENCE 348 AA; 37794 MW; 4D7B6D9089174049 CRC64;
Query Match 45.0%; Score 952.5; DB 1; Length 348;
Best Local Similarity 46.6%; Pred. No. 3.3e-62;
Matches 178; Conservative 57; Mismatches 106; Indels 41; Gaps 8;
OY 1 MSSRIARALAVLTLLHL-TRLAL-STCPACHCLE-APKAPGVGLVRDGGCCCKYCA 57
DB 1 MIAVAGPISALVYLLALCTPRATGQDSACQCAEAAPHPACVSLVDGCGCRCA 60
OY 58 KQLNEDCKPTDHTKGLCECNFGASSTALNGICRAQSGRRPCENSRITYNGESFPQNC 117
DB 61 KQGLGCTERPDPDHKGLFDFGSPANKRIGVCTAK-DGAPCVGSGSVSGESFQSSC 119
OY 118 OHQCTCIDGAVGCIPLCPQELSLPVLGCPNRLKAVYTGCCCEBWCDEDSIKDPREDQD 177
DB 120 KYQCTCIDGAVGCVPLCSMDVRLSPDCPFRRVRLPKCCBEMWCDEP----- 168
OY 178 LLGKELGSDASVELTRNNELIANGKSKRL-PVFGMEPRILYNPLQGGKCIYQTSW 236
DB 169 -----KDRIVGPAALAAVRLDETFSPDPMM-----RANCLVQTTW 205
OY 237 SOCSKTCGTGISTRTVNDNPECRILVKEIRICEVPRGCPVYSILKGRKCKSKTKSPBPV 296

```

Db      206  SACSCTGCGISRTVNDNTPCFLERKQSTLCWVRPCPEADLEENIKKGCIRTPKATXV 265
QY      297  RPTVACGLSVKRTPRKPCGSCVDGRCCTQTLRTVMRRCEDEFFSNVMMIOSCKN 356
Db      266  KKLSCCTGCTVKTFRFAFCVCTDGRCCPTHRTTLTVEFKCPDGEIKMKNMFIKTCACH 325
QY      357  YNCPHANEAAFPFY--RLFNDI 376
Db      326  YNCPGNDNIFESLYTKMTGDM 347

RESULT 7
CTGF HUMAN STANDARD; PRT; 349 AA.
ID      P29279; 0960X2;
AC      01-DEC-1992 (Rel. 24, Created)
DT      01-DEC-1992 (Rel. 24, Last sequence update)
DT      28-FEB-2003 (Rel. 41, Last annotation update)
DE      Connective tissue growth factor precursor (Hypertrophic chondrocyte-
DE      specific protein 24).
GN      CTGF OR HCS24.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX      NCBI_TaxID=9606;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      TISSUE=Umbilical vein endothelial cells;
RX      MEDLINE=93187114; PubMed=1293144;
RT      Igarashi A., Bradham D.M., Okochi H., Grotendorst G.R.;
RL      "Connective tissue growth factor.";
RL      J. Dermatol. 19:642-643(1992).
RN      [2]
RP      SEQUENCE FROM N.A.
RC      TISSUE=Umbilical vein endothelial cells;
RX      MEDLINE=93187114; PubMed=1293144;
RT      Igarashi A., Bradham D.M., Okochi H., Grotendorst G.R.;
RL      "Connective tissue growth factor.";
RL      J. Dermatol. 19:642-643(1992).
RN      [3]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=97207446; PubMed=9054739;
RA      Oemar B.S., Werner A., Garnier J.M., Do D.D., Godoy N., Nauck M.,
RA      Marz W., Rupp J., Pech M., Luescher T.F.;
RT      "Human connective tissue growth factor is expressed in advanced
RT      atherosclerotic lesions.";
RL      Circulation 95:831-839(1997).
RN      [4]
RP      SEQUENCE FROM N.A.
RA      Cobley V.;
RL      Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RN      [5]
RP      FUNCTION.
RC      TISSUE=Chondrocytes;
RX      MEDLINE=20080284; PubMed=10614647;
RA      Nakaiishi T., Nishida T., Shimo T., Kobayashi K., Kubo T.,
RA      Tamatani T., Tezuka K., Takigawa M.;
RT      "Effects of CTGF/HCS24, a product of a hypertrophic chondrocyte-
RT      specific gene, on the proliferation and differentiation of
RT      chondrocytes in culture.";
RL      Endocrinology 141:264-273(2000).
RN      [6]
RP      FUNCTION: MAJOR CONNECTIVE TISSUE MITOATTRACTANT SECRETED BY
RP      VASCULAR ENDOTHELIAL CELLS. PROMOTES PROLIFERATION AND
RP      DIFFERENTIATION OF CHONDROCYTES. MEDIATES CELL ADHESION AND
RP      ENHANCES FIBROBLAST GROWTH FACTOR-INDUCED DNA SYNTHESIS.
CC      -1- SUBUNIT: Monomer
CC      -1- SUBCELLULAR LOCATION: FOUND IN THE EXTRACELLULAR MATRIX AND AS A
CC      -1- ALTERNATIVE PRODUCTS:
CC      Event-Alternative splicing; Named isoforms-2;

```

```

CC      Name-Long;
CC      IsoId=P29279-1; Sequence=Displayed;
CC      Name-Short;
CC      IsoId=P29279-2; Sequence=VSP_002460;
CC      Note=No experimental confirmation available;
CC      -1- SIMILARITY: Contains 1 IGFBP domain.
CC      -1- SIMILARITY: Contains 1 WFPC domain.
CC      -1- SIMILARITY: Contains 1 TSP type-1 domain.
CC      -1- SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain.
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL: M92934; AAA91279.1; -
DR      EMBL: X78947; CAA55544.1; -
DR      EMBL: AL354866; CAC44023.1; -
DR      PIR: A40551; A40551.
DR      Genew; HGNC:2500; CTGF.
DR      MIM: 121009; -
DR      GO: GO:0005578; C:extracellular matrix; TAS.
DR      GO: GO:0005886; C:plasma membrane; TAS.
DR      GO: GO:0005520; F:insulin-like growth factor binding activity; TAS.
DR      GO: GO:0008151; P:cell growth and/or maintenance; TAS.
DR      GO: GO:0008544; P:epidermal differentiation; TAS.
DR      GO: GO:0009611; P:response to wounding; TAS.
DR      InterPro: IPR006208; Cys_knot.
DR      InterPro: IPR006207; Cys_knot.
DR      InterPro: IPR000867; Ins1_gro_fac_pr.
DR      InterPro: IPR000884; TSP1.
DR      InterPro: IPR001007; WFPC.
DR      Pfam: PF00007; Cys_knot; 1.
DR      Pfam: PF00219; IGFBP; 1.
DR      Pfam: PF00090; TSP_1; 1.
DR      Pfam: PF00093; vwc; 1.
DR      SMART: SM00041; CT; 1.
DR      SMART: SM00121; IB; 1.
DR      SMART: SM00209; TSP1; 1.
DR      SMART: SM00214; VWC; 1.
DR      PROSITE: PS01185; CTCK_1; 1.
DR      PROSITE: PS01225; CTCK_2; 1.
DR      PROSITE: PS00222; IGF_BINDING; 1.
DR      PROSITE: PS00092; TSP1; 1.
DR      PROSITE: PS01208; WFPC_1; 1.
DR      PROSITE: PS0184; WFPC_2; 1.
KW      Cell adhesion; DNA synthesis; Extracellular matrix; Signal;
KW      Alternative splicing.
FT      SIGNAL 1..26
FT      CHAIN 27..349
FT      DOMAIN 33..100
FT      DOMAIN 101..167
FT      DOMAIN 168..243
FT      DOMAIN 244..330
FT      DOMAIN 331..393
FT      DISULFID 256..283
FT      DISULFID 284..307
FT      DISULFID 308..335
FT      DISULFID 336..363
FT      CARBOHYD 28..28
FT      CARBOHYD 225..225
FT      CARBOHYD 226..226
FT      VARSPLIC 172
FT      CONFLICT 83..83
SQ      SEQUENCE 349 AA; 38069 MW; 0ECF8470B357EA95 CRC64;
Query Match 44.8%; Score 947; DB 1; Length 349;
Best Local Similarity 46.6%; Pred. No. 8; 3e-62;
Matches 174; Conservative 58; Mismatches 101; Indels 40; Gaps 8;

```


[illegible]

```

RESULT 8
CTGF_BOVIN
ID CTGF_BOVIN STANDARD: PRT: 349 AA.
AC 018739; Q8GJ71;
DT 15-JUL-1998 (Rel. 36, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Connective tissue growth factor precursor.
GN CTGF.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bos.
OX NCBI_TaxID:9913;
RN [1]
RP SEQUENCE FROM N.A.
RA TISSUE=Arteria;
RC Lillensiek B., Lin Z., Fortis T., Schimanski M., Biehans A.,
RA Kallitz M., Kaufmann G., Schweigerer L., Ziegler R., Nawroth P.P.;
RL Submitted (Aug-1997) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA TISSUE=Liver;
RC Matchans M., Schwitters C., Hove M., Rupp S., Erondou N.E.;
RA "Bovine connective tissue growth factor. organization of the
RT chromosomal gene and demonstration of promoter activity.";
RT Submitted (SEP-2000) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: MAJOR CONNECTIVE TISSUE MITOATTRACTANT SECRETED BY
CC VASCULAR ENDOTHELIAL CELLS. PROMOTES PROLIFERATION AND
CC DIFFERENTIATION OF CHONDROCYTES (BY SIMILARITY). MEDIATES CELL
CC ADHESION AND ENHANCES FIBROBLAST GROWTH FACTOR-INDUCED DNA
CC SYNTHESIS (BY SIMILARITY).
CC -1- SUBUNIT: Monomer (By similarity).
CC -1- SUBCELLULAR LOCATION: FOUND IN THE EXTRACELLULAR MATRIX AND AS A
CC SOLUBLE FORM (BY SIMILARITY).
CC -1- SIMILARITY: Contains 1 IGFBP domain.
CC -1- SIMILARITY: Contains 1 WPC domain.
CC -1- SIMILARITY: Contains 1 TSP type-1 domain.
CC -1- SIMILARITY: Contains 1 C-terminal cystine knot-like (CTCK) domain.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

[illegible]


```

Db 278 TYRANKGCVCTDRCCTPHRTTLPVPEKCPDSEWKKSMFKTCACHNCGDNDIFE 337
QY 368 PFY--RLFNDI 376
Db 338 SLTYRKMVGDM 348

RESULT 9
CTGF_PIG STANDARD; PRT; 349 AA.
ID CTGF_PIG
AC 019113;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Connective tissue growth factor precursor.
GN CTGF.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Uterus;
RX MEDLINE=97390475; PubMed=9242708;
RA Brigstock D.R., Steffen C.L., Kim G.Y., Vegunta R.K., Diehl J.R.,
RA Harding P.A.;
RT "Purification and characterization of novel heparin-binding growth
RT factors in uterine secretory fluids. Identification as heparin-
RT regulated Mr 10,000 forms of connective tissue growth factor.";
RL J. Biol. Chem. 272:20275-20282(1997).
CC -1- FUNCTION: MAJOR CONNECTIVE TISSUE MITOGENETIC SECRETED BY
CC VASCULAR ENDOTHELIAL CELLS. PROMOTES PROLIFERATION AND
CC DIFFERENTIATION OF CHONDROCYTES (BY SIMILARITY). MEDiates CELL
CC ADHESION AND ENHANCES FIBROBLAST GROWTH FACTOR-INDUCED DNA
CC SYNTHESIS (BY SIMILARITY).
CC -1- SUBUNIT: Monomer (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: FOUND IN THE EXTRACELLULAR MATRIX AND AS A
CC SOLUBLE FORM (BY SIMILARITY).
CC -1- SIMILARITY: Contains 1 IGFBP domain.
CC -1- SIMILARITY: Contains 1 WFEC domain.
CC -1- SIMILARITY: Contains 1 TSP type-1 domain.
CC -1- SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL, U83916; AAC48756.1.
CC InterPro: IPR006208; Cys_knot.
CC InterPro: IPR006207; Ins_gro_C.
CC InterPro: IPR000867; Ins_gro_fac_Pr.
CC InterPro: IPR000884; TSP1.
CC Pfam: PF00007; Cys_knot; 1.
CC Pfam: PF000219; IGFBP; 1.
CC Pfam: PF00093; TSP_1; 1.
CC SMART: SM00041; CT; 1.
CC SMART: SM00121; IB; 1.
CC SMART: SM00209; TSP1; 1.
CC SMART: SM00214; WVC; 1.
CC PROSITE: PS01185; CTCK_1; 1.
CC PROSITE: PS01225; CTCK_2; 1.
CC PROSITE: PS00232; IGF_BINDING; 1.
CC PROSITE: PS00982; TSP1; 1.
CC PROSITE: PS01208; WVC_1; 1.
CC PROSITE: PS0184; WVC_2; 1.
CC Cell adhesion; DNA synthesis; Extracellular matrix; Signal.

```

```

FT SIGNAL 1 26 POTENTIAL.
FT CHAIN 27 349 CONNECTIVE TISSUE GROWTH FACTOR.
FT DOMAIN 33 100 IGFBP.
FT DOMAIN 101 167 WFEC.
FT DOMAIN 198 243 TSP TYPE-1.
FT DOMAIN 256 330 CTCK.
FT DISULFID 256 293 BY SIMILARITY.
FT DISULFID 273 307 BY SIMILARITY.
FT DISULFID 284 323 BY SIMILARITY.
FT DISULFID 287 325 BY SIMILARITY.
FT DISULFID 292 329 BY SIMILARITY.
SQ SEQUENCE 349 AA; 38007 MW; BB510E2B2B52D4A0 CRC64;

Query Match 42.8%; Score 906; DB 1; Length 349;
Best Local Similarity 44.9%; Pred. No. 7.7e-59;
Matches 167; Conservative 57; Mismatches 108; Indels 40; Gaps 7;

QY 10 ALVYTLHLRLAL-STCPAACHCP-LEAPKCAPGVGLYRDGCGCCVCAKQINDECSKT 67
Db 12 AFVLLALCSRPASGDCSGGCCAGKRRACAGSYLVLDGCGCCRLCAKQIGELCTER 71
QY 68 QPCDHTKGLGECNFGASSIALKIGICRAOSEGRPCCEYSRIYONGESFQPMQHOCTCIDA 127
Db 72 DPCDPKRGJLFCDFGSPANKRIGVCTAK-DGAPCVFGTIVRSSEFSQSKYQCTCLDGA 130
QY 128 VGCIPICPELSPNCGENPRIVKYTGCCCEWVDEDSINDPMEDOGILGKELGFA 187
Db 131 VGVPLCSMDVRLPDCPPRRVKLPKGCCEWVDEP----- 169
QY 188 SEVELRNNEILLAVGGRSLRL-PVFGMEPRILYPLQKCIYVTTSMQSKTCGTG 246
Db 170 -----KDHVYVAGALAYRLIEDFFGDPFMM-----RANLVGTETMSACSKTCGNG 216
QY 247 ISTRVTNDPECRVKEPRICEVRPCGQPVYSLKGGKCKSKTKSPPEVRYAGCLSV 306
Db 217 ISTRVTNDNAFCFLKESKRLCWMRPEADLEENIKKGCIRTPKISKVKELSGCTSV 276
QY 307 KRYRPFYGCYDVRGCTPQLRTYVMKRPCEGEFFSNVMMIOGCKNYPHANEAA 366
Db 277 KYTRAFGCVCTDRCCTPHRTTLPVPEKCPDSEWKKSMFKTCACHNCGDNDIF 336
QY 367 PFY--RLFNDI 376
Db 337 ESLYRKMVGDM 348

RESULT 10
NOV_CHICK STANDARD; PRT; 351 AA.
ID NOV_CHICK
AC P28686;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE NOV protein precursor (Nephroblastoma overexpressed gene protein).
GN NOV.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Brown leghorn;
RX MEDLINE=92107157; PubMed=1309586;
RA Joliet V., Martinerie C., Dambrine G., Plassiat G., Brissac M.,
RA Crochet J., Perbal B.;
RT "Proximal rearrangements and overexpression of a new cellular gene
RT (nov) in myeloblastosis-associated virus type 1-induced
RT nephroblastomas";
RL Mol. Cell. Biol. 12:10-21(1992).
CC -1- FUNCTION: IMMEDIATE-EARLY PROTEIN LIKELY TO PLAY A ROLE IN CELL
CC GROWTH REGULATION. ITS OVEREXPRESSION IS ASSOCIATED WITH
CC TUMORIGENESIS AND EXPRESSION OF A N-TERMINAL-TRUNCATED VERSION

```

CC OF NOV GENE IN CHICKEN EMBRYONIC FIBROBLASTS (CEF) IS SUFFICIENT
 CC TO INDUCE THE TRANSFORMATION OF CEF IN VITRO.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: BRAIN AND HEART, AND AT A LOWER LEVEL IN
 CC MUSCLE AND INTESTINE, IN THE EMBRYO, LUNG AND LESS SO IN BRAIN AND
 CC SPLEEN, IN ADULT CHICKEN.
 CC -1- DEVELOPMENTAL STAGE: MAWI-INDUCED NEPHROBLASTOMAS EXPRESS A HIGH
 CC LEVEL OF NOV GENE WHOSE TRANSCRIPTION IS NORMALLY ARRESTED IN
 CC ADULT KIDNEY.
 CC -1- SIMILARITY: Contains 1 IGFBP domain.
 CC -1- SIMILARITY: Contains 1 WFMC domain.
 CC -1- SIMILARITY: Contains 1 TSP type-1 domain.
 CC -1- SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X59284; CAA41975.1; -
 DR PIR: S20078; S20078.
 DR InterPro: IPR006208; Cys_knot.
 DR InterPro: IPR006207; Cys_knot_C.
 DR InterPro: IPR000867; Ins1_gro_fac-pr.
 DR InterPro: IPR000884; TSP1.
 DR InterPro: IPR001007; WFMC_C.
 DR Pfam: PF00007; Cys_knot.1.
 DR Pfam: PF00219; IGFBP.1.
 DR Pfam: PF00090; TSP1.1.
 DR Pfam: PF00093; WVC.1.
 DR Pfam: PF00041; CT.1.
 DR SMART: SM00121; IB.1.
 DR SMART: SM00209; TSP1.1.
 DR SMART: SM00214; WVC.1.
 DR PROSITE: PS01185; CTCK_1; 1.
 DR PROSITE: PS01225; CTCK_2; 1.
 DR PROSITE: PS00222; IGF_BINDING; 1.
 DR PROSITE: PS50092; TSP1.1.
 DR PROSITE: PS01208; WFMC.1; 1.
 DR PROSITE: PS50184; WFMC_2; 1.
 DR Proto-oncogene; Growth factor; Signal.
 KW SIGNAL 1 24
 FT CHAIN 25 351 POTENTIAL.
 FT DOMAIN 31 103 NOV PROTEIN.
 FT DOMAIN 104 170 IGFBP.
 FT DOMAIN 201 246 WFMC.
 FT DOMAIN 258 332 CTCK.
 FT DISULFID 258 295 BY SIMILARITY.
 FT DISULFID 275 309 BY SIMILARITY.
 FT DISULFID 286 325 BY SIMILARITY.
 FT DISULFID 289 327 BY SIMILARITY.
 FT DISULFID 294 331 BY SIMILARITY.
 FT CARBOHYD 274 274 N-LINKED (GLCNAc. . .) (POTENTIAL).
 SQ SEQUENCE 351 AA; 38268 MW; 1ECB3FA30586C797 CRC64;
 Query Match 40.2%; Score 851.5; DB 1; Length 351;
 Best Local Similarity 44.0%; Pred. No. 6,9e-55;
 Matches 164; Conservative 44; Mismatches 104; Indels 61; Gaps 8;
 QY 11 LVVTLHLRLALS---TCPAAC--HCPLEAPKAPGVGLVYRDGCGCCCKCAKOLNEDC 64
 DB 12 LLLLLLLRPCEVSGREAACPCCGRCPPAEPFPAVLDGCGCLVCARORGSC 71
 QY 65 STQPCDHRTKGLKCFNGASSTALKICRAOSGRCEVNSRYONGESFQPCQHCQCI 124
 DB 72 SPLPCDSDSGGLYCRGPDGEGGAGICMW-LEGDMCVDGMVYRNGETFPQCKYQCTCR 130
 QY 125 DGAAGCIPICQPELSPNIGCPNPLRYVTGGCCCEWYCDDESDISDPEDDQGLLG---- 180
 DB 131 DQIIGCLPRCNIGLILPDPDCEFPKRIEYVGGCCCKRWVC-----DPRDEV--LLGGRAM 182

QY 181 -----KELGPDASEVELTRNNELIANGKGRSLKRLPYFGMERILLNPLOQOKIYQTT 234
 DB 183 AAYROEATLIDVSD-----SSANCIQQT 207
 QY 235 SMSGCKTGCGTGISTRYTNDNPECRLYKETRICVPRPGQPVYSLKGRKCKSKTSKPE 294
 DB 208 EWSACSCKSGMGFSTRYTNNOQCEMYKQTRLCMRCEDEPSD-KKGRKCIQTKSMK 266
 QY 295 PVRETYAGCLSVKKYRKYRGSCVDRCCPDLTRYKMRFRDEGETFSKNVMIQSCK 354
 DB 267 AVRETYKNCISVYQYKRYRGCLGNDRCCTPHNTKITQVERPCQGFLLKPMMLINTCV 326
 QY 355 CNINCPHANEAF 367
 DB 327 CHGNCPOSNNAF 339
 RESULT 11
 NOV_COTJA STANDARD; PRT; 353 AA.
 AC P42642;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE NOV protein precursor (Nephroblastoma overexpressed gene protein).
 GN NOV.
 OS Coturnix coturnix japonica (Japanese quail).
 CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 CC Coturnix
 CC NCBI_Taxid=93934;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Weiskirchen R., Bister K.;
 RL Submitted (AUG-1994) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: IMMEDIATE-EARLY PROTEIN LIKELY TO PLAY A ROLE IN CELL
 CC GROWTH REGULATION (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Secreted (By similarity).
 CC -1- SIMILARITY: Contains 1 IGFBP domain.
 CC -1- SIMILARITY: Contains 1 WFMC domain.
 CC -1- SIMILARITY: Contains 1 TSP type-1 domain.
 CC -1- SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U13063; AAA21128.1; -
 DR InterPro: IPR006208; Cys_knot.
 DR InterPro: IPR006207; Cys_knot_C.
 DR InterPro: IPR000867; Ins1_gro_fac-pr.
 DR InterPro: IPR000884; TSP1.
 DR InterPro: IPR001007; WFMC_C.
 DR Pfam: PF00007; Cys_knot.1.
 DR Pfam: PF00219; IGFBP.1.
 DR Pfam: PF00090; TSP1.1.
 DR Pfam: PF00093; WVC.1.
 DR SMART: SM00041; CT.1.
 DR SMART: SM00121; IB.1.
 DR SMART: SM00209; TSP1.1.
 DR SMART: SM00214; WVC.1.
 DR PROSITE: PS01185; CTCK_1; 1.
 DR PROSITE: PS01225; CTCK_2; 1.
 DR PROSITE: PS00222; IGF_BINDING; 1.
 DR PROSITE: PS50092; TSP1.1.
 DR PROSITE: PS01208; WFMC.1; 1.
 DR PROSITE: PS50184; WFMC_2; 1.
 KW Proto-oncogene; Growth factor; Signal.

```

FT SIGNAL 1 26 POTENTIAL.
FT CHAIN 27 353 NOV PROTEIN.
FT DOMAIN 33 105 IGFBP.
FT DOMAIN 106 172 VMRC.
FT DOMAIN 203 248 TSP TYPE-1.
FT DOMAIN 260 334 CTCK.
FT DISULFID 260 297 BY SIMILARITY.
FT DISULFID 260 311 BY SIMILARITY.
FT DISULFID 277 311 BY SIMILARITY.
FT DISULFID 288 327 BY SIMILARITY.
FT DISULFID 291 329 BY SIMILARITY.
FT DISULFID 296 333 BY SIMILARITY.
FT CARBOHYD 276 276 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 353 AA; 38667 MW; 717D9F853382E89 CRC64;

Query Match 40.1%; Score 848; DB 1; Length 353;
Best Local Similarity 43.3%; Pred. No. 1,2e-54;
Matches 164; Conservative 45; Mismatches 104; Indels 66; Gaps 8;

QY 10 ALVVTLLHTRAL-----STCPAAC--HCPLPAKCAPGVLPDCCGCKVKAR 58
DB 8 SLVYVLLLLLLLLLRLRESEVNGREARPCRCGRCRPAEPRRCAGVRAVDGCGCLYCAR 67
QY 59 QLNEDCSKTPCCHTRKGLCECNFGASSTALGICRAOSEGRPCRYNSRTYQNGSPQNCQ 118
DB 68 QRESECSPLLPCEDESGGLYCDRGPEDEGGGTGICMV--LEGDNVCFDMYRNGETFQPSCK 126
QY 119 HOCTCIDGAVGCIPLCEQELSLNLCGPNRLKVTGQCCCEWVCDSDTKDMEDODGL 178
DB 127 YQCTCRDQIGCLPRCLMLGLLPDPCPEPRKTEVQECCCKVVC-----EPREDEV--L 178
QY 179 LG-----KELGFDASEVELTRNNELIAVGKGRSLRLVFGMEPRILYNPLOGQR 228
DB 179 LGGFAMAAYRQENATLGDVSD-----SSAN 203
QY 229 CIYQTTSMQSCSTCTGCTGISTRYTNDNPECRLYKERRICVRPGOGPVYSLKKGKCK 288
DB 204 CIOQTETMSACSSCGSGFSTRVTNRNOCEWVKOTRLCMRRCENEEPPSD--KKGKCIIR 262
QY 289 TKKSPREFTYAGCLSVKRYRPGYSCVDGRCPCPOLRRTYKMRPRCDGFTFSNVM 348
DB 263 TKKSMANVREYNCTSVQYTKRFGCLGMDGRCCTPHNTKTIQVERFCQGFLLKPKPM 322
QY 349 MIQSCKCNVNCPRHANEAF 367
DB 323 LINTCVCHGNCQPSNNAFF 341

RESULT 12
NOV_HUMAN STANDARD; PRT; 357 AA.
ID NOV_HUMAN
AC P48745; 096BY5;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE NOV protein homolog precursor (NOVH) (Nephroblastoma overexpressed
  gene protein homolog).
GN NOV.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=94336229; PubMed=7520150;
RA Martinie C., Huif V., Joubert I., Badzioch M., Saunders G.F.,
RA Strong L.C., Perbal B.;
RT "Structural analysis of the human nov proto-oncogene and expression
  in Wilms tumor";
RT Oncogene 9:2729-2732(1994).
RN [2]
RP SEQUENCE FROM N.A.
RA Jiang D., Gou D., Li W.;

```

```

RT "Cloning, sequencing and expression of human nov gene.";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.L., Wang J., Hsieh F.,
RA Diatchenko L., Marusla K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McKernan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.D., Huiyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smilus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
  human and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [4]
RP INTERACTION WITH FBLN1.
RX PubMed=9927660;
RA Perbal B., Martinie C., Salason R., Werner M., He B., Rolzman B.;
RT "The C-terminal domain of the regulatory protein NOVH is sufficient to
  promote interaction with fibulin 1C: a clue for a role of NOVH in
  cell-adhesion signaling.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:869-874(1999).
RT
CC - FUNCTION: IMMEDIATE-EARLY PROTEIN LIKELY TO PLAY A ROLE IN CELL
  GROWTH REGULATION (BY SIMILARITY).
CC - SUBUNIT: Interacts with FBLN1.
CC - SUBCELLULAR LOCATION: Secreted.
CC - TISSUE SPECIFICITY: INCREASED EXPRESSION IN WILMS TUMOR OF THE
  STROMAL TYPE.
CC - SIMILARITY: Contains 1 IGFBP domain.
CC - SIMILARITY: Contains 1 VMRC domain.
CC - SIMILARITY: Contains 1 TSP type-1 domain.
CC - SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement. (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sid.ch.)
CC
DR EMBL: X78351; CAAS5146.1; -
DR EMBL: X78352; CAAS5146.1; JOINED.
DR EMBL: X78353; CAAS5146.1; JOINED.
DR EMBL: X78354; CAAS5146.1; JOINED.
DR EMBL: X96584; CAAS5403.1; -
DR EMBL: AY082381; AAL92490.1; -
DR EMBL: BC015028; AAL15028.1; -
DR PIR: I38069; I38069.
DR Genew: HGNC:7885; NOV.
DR MIM: 164958; -
DR InterPro: IPR006208; Cys_knot.
DR InterPro: IPR006207; Cys_knot_C.
DR InterPro: IPR000867; Insi_gro_fac_pr.
DR InterPro: IPR000884; TSP1.
DR InterPro: IPR001007; VMF_C.
DR Pfam: PF00007; Cys_knot; 1.
DR Pfam: PF00219; IGFBP; 1.
DR Pfam: PF00090; TSP_1; 1.
DR Pfam: PF00093; VMF; 1.
DR SMART: SM00041; CT; 1.

```

DR	SMART:	SM00121;	IB; 1.
DR	SMART:	SM00209;	TSP1; 1.
DR	SMART:	SM00214;	VWC; 1.
DR	PROSITE;	PS01185;	CTCK_1; 1.
DR	PROSITE;	PS01285;	CTCK_2; 1.
DR	PROSITE;	PS00222;	IGF_BINDING; 1.
DR	PROSITE;	PS50092;	TSP1; 1.
DR	PROSITE;	PS01206;	VWFC_1; 1.
DR	PROSITE;	PS0184;	VWFC_2; 1.
KW	Growth factor;	Signal.	
FT	SIGNAL	1	27
FT	CHAIN	28	357
FT	DOMAIN	35	107
FT	DOMAIN	108	174
FT	DOMAIN	205	250
FT	DOMAIN	264	338
FT	DISULFID	264	301
FT	DISULFID	281	315
FT	DISULFID	292	331
FT	DISULFID	295	333
FT	DISULFID	300	337
FT	CARBOHYD	97	97
FT	CARBOHYD	280	280
FT	CONFLICT	97	97
SO	SEQUENCE	357 AA;	39162 MW; 035D5BF4576BD85B CRR64;

Query Match	39.1%;	Score	827.5;	DB	1;	Length	357;
Best Local Similarity	43.1%;	Pred. No.	3.8e-53;				
Matches	157;	Conservative	48;	Mismatches	120;	Indels	39;
						Gaps	7

Oy	7	RALILVLLIHLHRLTALST--CPAAC--HCPLAARCAPAGVLVRDGGCCVCARQLNE	62
Dd	14	QCICITFLILLHLGAVATQRCPPCPCRCRPAITPTCAAGVAIVLDGSCCLVCARQRE	73
Oy	63	DSCKTPQCDHTKGLCECNFGASSTALKIGICRAOSERPCEYNSRIYQNGESFPNCQHCT	122
Dd	74	SCSDLEPDESSGLICDRSADPSNQTGICTA--VEGDNCVFDFVIYIRSGEKFOPSKFOCT	132
Oy	123	CIDGAVGCIPICPQBELSLPNLGCAPPRILYKVAVGCCCEWVCBDESIKDPMEOQDLIGE	182
Dd	133	CRDGIQGVCPKQIDLVLPEPNCAPRKVEYVGECCERMKIGPD-----EDSIGILT	185
Oy	183	LGFDASEVELTRNNELIAVGKGRSLKLRFVEGMEIRILIYNPLQGOKCIAQVTTSWSCSKT	242
Dd	186	LAAYREPETL-----CYEVSD-----SVNCHIQTEWTAGCSKS	219
Oy	243	CGTGISTRYTNDNPBCRLVKETRICEVARPCGO--PYVSLLKKGKKCSKTUKSPDYAFETYA	301
Dd	220	CGMGSTVTYTNNRQCEMLKQTRLCMVRRCEDEPEQPTDKKGKKCLRTKKSILKAHLQRC	279
Oy	302	GCLSYAKTKRPKYCGSCVDGRCTPOLTRTYVKRRFRCDEGETSKVMYMIOQSKCANCPH	361
Dd	280	NCTSLHYTKPRFCGCVSGDRCTPHNTKTIQAEOFCSPOAIYKKPVMVIGTCCTCHNCPR	339
Oy	362	ANEA 365	
Dd	340	NNEA 343	

RESULT 13			
NOV_XENILA			
ID	NOV_XENILA	STANDARD:	PRT: 343 AA.
AC	P51609:		
DT	01-OCT-1996 (Rel. 34, Created)		
DT	01-OCT-1996 (Rel. 34, Last sequence update)		
DT	28-FEB-2003 (Rel. 41, Last annotation update)		
DE	NOV protein homolog precursor (Xnov).		
GN	NOV.		
OS	Xenopus laevis (African clawed frog).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;		
OC	Xenopodinae; Xenopus.		
OX	NCBI_TaxID=8355;		

```

RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=96257227; PubMed=8666280;
RA      Ying Z., King M.L.;
RT      "Isolation and characterization of xnov, a Xenopus laevis ortholog of
RT      the chicken nov gene.";
RL      Gene 171:243-248(1996).
CC      -1- FUNCTION: IMMEDIATE-EARLY PROTEIN LIKELY TO PLAY A ROLE IN CELL
CC      GROWTH REGULATION (BY SIMILARITY).
CC      -1- SUBCELLULAR LOCATION: Secreted.
CC      -1- SIMILARITY: Contains 1 IGBP domain.
CC      -1- SIMILARITY: Contains 1 WPC domain.
CC      -1- SIMILARITY: Contains 1 TSP type-1 domain.
CC      -1- SIMILARITY: Contains 1 C-terminal cystine knot-like (CTCK) domain.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation
CC      at the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).

```

DR	EMBL:	U37063;	AAB17096.1;	-.
DR	InterPro:	IPR006208;	Cys_knot.	
DR	InterPro:	IPR006207;	Cys_knot.C.	
DR	InterPro:	IPR000867;	InsL_gro_fac.pr.	
DR	InterPro:	IPR000884;	TSP1.	
DR	InterPro:	IPR001007;	vWF_C.	
DR	pfam:	PF000007;	Cys_knot_1.	
DR	pfam:	PF00219;	IGFBP_1.	
DR	pfam:	PF00090;	tsp_1; 1.	
DR	pfam:	PF00093;	wvc; 1.	
DR	SMART;	SM00041;	CT; 1.	
DR	SMART;	SM00121;	IB; 1.	
DR	SMART;	SM00209;	TSP1; 1.	
DR	SMART;	SM00214;	WVC; 1.	
DR	PROSITE:	PS01185;	CTCK_1; FALSE_NEG.	
DR	PROSITE:	PS01225;	CTCK_2; 1.	
DR	PROSITE:	PS00222;	IGF_BINDING; 1.	
DR	PROSITE:	PS50092;	TSP1; 1.	
DR	PROSITE:	PS01208;	WVFC_1; 1.	
DR	PROSITE:	PS50184;	WVFC_2; 1.	
KM	Growth factor:	Signal.		
FT	SIGNAL	1	18	POTENTIAL.
FT	CHAIN	19	343	NOV PROTEIN HOMOLOG.
FT	DOMAIN	21	92	IGFBP.
FT	DOMAIN	93	159	WVFC.
FT	DOMAIN	190	235	TSP TYPE-1.
FT	DOMAIN	249	323	CTCK.
FT	DISULFID	249	286	BY SIMILARITY.
FT	DISULFID	266	300	BY SIMILARITY.
FT	DISULFID	277	316	BY SIMILARITY.
FT	DISULFID	280	318	BY SIMILARITY.
FT	DISULFID	285	322	BY SIMILARITY.
FT	CAROHWD	265	265	N-LINKED (GICNAC..)(POTENTIAL).
SQ	SEQUENCE	343 AA;	38070 MW;	677D7078EBE21365F CXC64;

Query Match	39.0%;	Score	825.5;	DB 1;	Length	343;
Best Local Similarity	41.8%;	Pred. No.	5,2e-53;			
Matches	158;	Conservative	44;	Mismatches	113;	Indels
					63;	Gaps
					9;	
QY	9	LALVYTLHLRLALSTPAAAC-HCPLEAPKCAPEVGLVRBOGCGCKKCAQALNMDCKRT	67			
DB	5	LALCFILL-100VASQKCPSCDDCPPEPPSCAPSVLLILDGCGCPYCARQEBESCSHL	63			
QY	68	QPCDHITKLECNFGASSTALKGICGAGSEGRPCENSRITYONGSEFQNCOHCTCIGA	127			
DB	64	NPQQRDKGYLCEFNDDPMEETGTCA-LEGNSCVFDGAVYRRRRESFQSCRYHCTCLNGH	122			
QY	128	VGCIPLCFQELSLPMIAGCPNRLVAVYTGQCCSEEWCDSDSKIDPMEDDGLG-----	180			
DB	123	IGCVFRCNDLLLPDPDCEPRPRRVAVPECECEKWCDSD---KEEN-----AIGGRMAAY	174			

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 5, 2003, 14:02:31 ; Search time 33.0868 Seconds

(without alignments)
2971.515 Million cell updates/sec

Title: US-09-495-448a-4
Perfect score: 2116
Sequence: 1 MSSRIARALALVYTLHLTR.....ANEAPFFYRLFDIHFRD 381

Scoring table:
BIOSUM62
Gapop 10.0 , Gapext 0.5*

Searched: 830525 seqs, 258052604 residues
Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_23:*
1: sp.archaea:*
2: sp.bacteria:*
3: sp.fungi:*
4: sp.human:*
5: sp.invertebrate:*
6: sp.mammal:*
7: sp.mhc:*
8: sp.organelle:*
9: sp.phage:*
10: sp.plant:*
11: sp.rodent:*
12: sp.virus:*
13: sp.vertebrate:*
14: sp.unclassified:*
15: sp.virus:*
16: sp.bacteriap:*
17: sp.archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1928	91.1	379	11	Q9WTM9
2	1663.5	78.6	375	13	Q98TX5
3	1587	75.0	334	4	Q9UID7
4	938	44.3	343	13	Q42607
5	930	44.0	344	13	Q98T08
6	912.5	43.1	347	13	Q9PT80
7	878	41.5	349	6	Q97765
8	822	38.8	354	11	Q8CA67
9	767	36.2	367	11	Q95388
10	758	35.8	367	11	Q99PP0
11	758	35.8	367	11	Q54775
12	626.5	29.6	331	4	Q95958
13	613	29.0	354	4	Q95389
14	514.5	24.3	251	11	Q8CIC8
15	495.5	23.4	280	4	Q9HCS3
16	485.5	22.9	176	13	Q9PS56

17	373.5	17.7	128	11	Q9R2C0	Q9R2C0
18	341	16.1	113	11	Q9Z164	Q9Z164
19	337	15.9	100	4	Q9UD16	Q9UD16
20	336	15.9	119	11	Q91V29	Q91V29
21	336	15.9	119	11	Q920W6	Q920W6
22	302	14.3	374	5	Q9VVK0	Q9VVK0
23	294	13.9	470	5	Q9VVK3	Q9VVK3
24	293	13.8	230	4	Q8WYK7	Q8WYK7
25	195	9.2	77	4	Q9UDE4	Q9UDE4
26	187	8.8	2327	13	Q91BG7	Q91BG7
27	176	8.3	70	13	Q9DF21	Q9DF21
28	174	8.2	1045	5	Q8T3A6	Q8T3A6
29	174	8.2	1070	5	Q8T3A7	Q8T3A7
30	174	8.2	1111	5	Q9XMD6	Q9XMD6
31	170.5	8.1	1036	4	Q9NZV1	Q9NZV1
32	169	8.0	1664	5	Q9TVQ2	Q9TVQ2
33	169	8.0	1704	5	Q94446	Q94446
34	168.5	8.0	1028	11	Q9JLL0	Q9JLL0
35	166.5	7.9	1048	13	Q8AWW5	Q8AWW5
36	164	7.8	58	6	Q97574	Q97574
37	162.5	7.7	4998	11	Q8CG65	Q8CG65
38	161.5	7.6	2104	5	Q21281	Q21281
39	161.5	7.6	2104	5	Q964N4	Q964N4
40	161	7.6	360	5	Q8MN64	Q8MN64
41	159.5	7.5	406	11	Q92513	Q92513
42	159.5	7.5	426	11	Q8VZA6	Q8VZA6
43	159	7.5	5146	6	Q8SPW4	Q8SPW4
44	158.5	7.5	792	13	Q90Z43	Q90Z43
45	158.5	7.5	1095	13	Q90XG4	Q90XG4

ALIGNMENTS

RESULT 1	ID	Q9WTM9	PRELIMINARY:	PRT:	379 AA.
AC	Q9WTM9	01-NOV-1999 (TREMBLrel. 12, Created)			
DT	01-NOV-1999 (TREMBLrel. 12, Last sequence update)				
DT	01-MAR-2003 (TREMBLrel. 23, Last annotation update)				
DE	CYR61 precursor.				
GN	CYR61.				
OS	Rattus norvegicus (Rat).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.				
OX	NCBI_Taxid=10116;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=Ism; TISSUE=Aorta;				
RA	Unoki H., Yonekura H., Furukawa K., Yamamoto H.;				
RT	"Rat CyR61 mRNA."				
RL	Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.				
DR	EMBL: AB015877; BAA78339.1; -				
DR	InterPro: IPR006208; Cys_knot.				
DR	InterPro: IPR006207; Cys_knot_C.				
DR	InterPro: IPR000867; Insl_gro_fac_pr.				
DR	InterPro: IPR000884; TSP1.				
DR	InterPro: IPR001007; VWF_C.				
DR	Pfam: PF00007; Cys_knot; 1.				
DR	Pfam: PF00219; IGFBP; 1.				
DR	Pfam: PF00090; TSP_1; 1.				
DR	Pfam: PF00093; VWC; 1.				
DR	SMART: SM00041; CT; 1.				
DR	SMART: SM00209; TSP1; 1.				
DR	SMART: SM00214; VWC; 1.				
DR	PROSITE: PS01185; CTCK_1; 1.				
DR	PROSITE: PS01225; CTCK_2; 1.				
DR	PROSITE: PS00222; IGF_BINDING; 1.				
DR	PROSITE: PS00092; TSP1; 1.				
DR	PROSITE: PS01208; VWF_C; 1.				
KW	Signal.				

FT	SIGNAL	1	24	POTENTIAL.
FT	CHAIN	25	379	CYR61.
SO	SEQUENCE	379 AA:	41728 MW:	D2ABAD77B84762B CRC64:
	Query Match		91.1%:	Score 1928; DB 11; Length 379;
	Best Local Similarity		90.6%:	Pred. No. 2,8e-179;
	Matches 347;	Conservative 11;	Mismatches 19;	Indels 6; Gaps 2
QY	1	MSSSIARALAVLWTLHTLRLALSTCPAACHCPLEAPCAPGVGLVPGCGCCKVCARQL	60	
DB	1	MSSTIKTLAAVATLHLTRLALSTCPASCHCPLEAPCAPGVGLVPGCGCCKVCARQL	60	
QY	61	NEDCSKTQPCDHTGTGLCECNFGASSTALGICRAOSEGRPCXEYSRIYONGESFOPNCQH	120	
DB	61	NEDCSKTQPCDHTGTGLCECNFGANSPTALGICRAOSEGRPCXEYSRIYONGESFOPNCQH	120	
QY	121	CTCIDGAAGCIPLCPOEILSNLGCNPRLKYVNTQCCCEEWCCDESDISTKDPNEDDGLG	180	
DB	121	CTCIDGAAGCIPLCPOEILSNLGCNPRLKYVNTQCCCEEWCCDESDISTKDPNEDDGLG	180	
QY	181	KELGFDASEVELTRNNELIANGKGRSLRLPVFGMEPRILYNPL--OGQKCIYQTTSMQ	238	
DB	179	--LGFDASEVELTRNNELIANGKSSLRRLPVFGMEPRILYNPLAHGQKCIYQTTSMQ	236	
QY	239	CSKTCIGTISTRYVNDNPECRILYKTRICEVAPCGQPYYSLLKGGKCKSKTKSPPEVF	298	
DB	237	CSKSGCTGISTRYVNDNPECRILYKTRICEVAPCGQPYYSLLKGGKCKSKTKSPPEVF	296	
QY	299	TYAGLSTKVRKPRKYGSCGVGRCTPDLTFRVKKRFCEGEFTSKNNVMQOSCKN	358	
DB	297	TYAGLSTKVRKPRKYGSCGVGRCTPDLTFRVKKRFCEGEFTSKNNVMQOSCKN	356	
QY	359	CPHANEAFPPYRLFNDIHKFRD	381	
DB	357	CPHNEASFRLYSLFNDIHKFRD	379	
RESULT 2				
ID	Q98TX5	PRELIMINARY:	PRT:	375 AA.
AC	Q98TX5:			
DT	01-JUN-2001 (Tremblrel. 17, Created)			
DT	01-JUN-2001 (Tremblrel. 17, Last sequence update)			
DT	01-MAR-2003 (Tremblrel. 23, Last annotation update)			
DE	Secreted cysteine-rich protein cyr61.			
OS	Xenopus laevis (African clawed frog).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;			
OC	Xenopodinae; Xenopus.			
NC	NCBI_TaxID=8355;			
RP	SEQUENCE FROM N.A.			
RA	LaLinkic B.V., Bennett B., Smith J.C.;			
RT	"Characterization of Xenopus cyr61."			
RL	Submitted (NOV-2000) to the EMBL/Genbank/DBJ databases.			
DR	EMBL: AF320592; AAK0947.1; -			
DR	InterPro: IPR006208; Cys_knot.			
DR	InterPro: IPR006207; Cys_knot_C.			
DR	InterPro: IPR00867; Insl_gro_fac.pr.			
DR	InterPro: IPR001007; WFE_C.			
DR	Pfam: PF00007; Cys_knot; 1.			
DR	Pfam: PF00090; tsp_1; 1.			
DR	Pfam: PF00093; vwc; 1.			
DR	SMART: SM00041; CT; 1.			
DR	SMART: SM00121; IB; 1.			
DR	SMART: SM00209; TSP1; 1.			
DR	SMART: SM00214; VWC; 1.			
DR	PROSITE: PS01185; CTCK_1; 1.			
DR	PROSITE: PS01225; CTCK_2; 1.			
DR	PROSITE: PS00222; IGF_BINDING; 1.			
DR	PROSITE: PS00922; TSP1; 1.			

DR	PROSITE: PS01208: YMF6.1	78.6%;	Score 1663.5;	DB 13;	Length 375;
SQ	SEQUENCE	375 AA;	41460 MW;	78075CA7B380304E CRC64;	
	Query Match	78.6%;	Score 1663.5;	DB 13;	Length 375;
	Best Local Similarity	79.3%;	Pred. No. 1.5e-153;		
	Matches	295;	Conservative	29;	Mismatches 41;
				Indels	7;
				Gaps	5
QY	11	LVYTLIILH-LTRALSTCPAACHCPLFEAKRCAGVGLYRDGCGCCCAVCYAKQLNEDGSKTP	69		
Db	10	LAIALSGFIDLAVSSCPAVCCPEYKRCAGVGLYVDGCGCCCKICAKQLNEDGSKTHP	69		
QY	70	CDHTKGLGECNFGASSTALKIGICRAOESGRPCENSRITONGSEFPNCOHOCTCIDGAWG	129		
Db	70	CDHTKGLGECNFGASSTALKIGICRAOESGRPCENSRITONGSEFPNCKHOCTCIDGAWG	129		
QY	130	CIPICPQELSLPNLGCPCNDRIVKYVTGQCCCEWVCDSDIKDPMEDQDGLGKELGPDASE	189		
Db	130	CLPILCPQELSLPNLGCPCNDRIVKYVGGCCCEWVCDSE--AKDPEVDMADDFENKEFGMDTNE	187		
QY	190	VELTFENNELIANGKRSRLRLPVPMEPRILLNPLNLOGKCIYQUTSWSCGSGTCGTGIST	249		
Db	188	GLTRFKNEPFAVIKQ-GLKMLPEVFESDDQ--SHVENSRCIYQUTSWSCGSGTCGTGIST	244		
QY	250	RYTNDNPCECRUYKEFRICEVVRPCGQPVYSSLLKGGKSKCTKKSPEPVFTVYAGCLSSVKRY	309		
Db	245	RYTNDNSNCRUYKEFRICEVVRPCGQPSVTSLLKGGKCKTKTKSQAPVRTTYAGCSSVKRY	304		
QY	310	RPKYCGSCVVDGRCCTPOLRTIVYKMFRCEDGETFSKNVMMIIOGSKCNVNCNPHANEAAPFC	369		
Db	305	RPKYCGSCVVDGRCCTPOOTRTIVKIFRCEDEGETFTKKNVMMIIOGSKCNVNCNPHATNE-APPY	363		
QY	370	YRLFNDIHKFRD	381		
Db	364	YRLFNDIHKFRD	375		

```

RESULT 3
Q9UID7
ID Q9UID7 PRELIMINARY; PRT; 334 AA.
AC Q9UID7;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE CyRel protein.
GN CyRel.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_Taxid=9606;
[1]
RN SEQUENCE OF 1-107 FROM N.A.
RC TISSUE=Kidney;
RA Anding B., Long Y.;
RT "Cloning of a new gene down-regulated in the small-cell tumor
RL submylonal-rhabdomyosarcoma (RMS).";
EMBL; AF003114; AAF21597.1; -
DR InterPro: IPR006208; Cys_knot.
DR InterPro: IPR006207; Cys_knot.C.
DR InterPro: IPR000867; Insl_gro_fac.pr.
DR InterPro: IPR000884; TSPL.
DR InterPro: IPR001007; vwf.C.
DR Pfam: PF00007; Cys_knot. 1.
DR Pfam: PF00219; IGFBP. 1.
DR Pfam: PF00090; tsp_1. 1.
DR Pfam: PF00093; wvc. 1.
DR SMART: SM0041; CT. 1.
DR SMART: SM00209; TSPL. 1.
DR SMART: SM00214; VMC. 1.
DR PROSITE: PS01225; CTCK_2. 1.
DR PROSITE: PS50092; TSPL. 1.
DR PROSITE: PS01208; WVCFC. 1.
SEQUENCE 334 AA: 37246 MW: 91898987A7352E948 CRC64:

```


Query Match 75.0%; Score 1587; DB 4; Length 334;
 Best Local Similarity 85.8%; Pred. No. 3.8e-146;
 Matches 295; Conservative 5; Mismatches 22; Indels 22; Gaps 4;

OY 49 GCGCCKVCACQALNEDCSKTQPODHTKLECNFGASSTALKICRAOSEGRPCENSRITQ 108
 DB 2 GCGTHPNCICHTIGHASTPTSKYKHHKLECNFGASSTALKICRAOSEGRPCENSRITQ 61
 OY 109 NGESQPNQCHQCTCIDAVAGCIPCLPQELSLPNIGCNPRILYKVTGGCCCEWVCDDESI 168
 DB 62 NGESQPNQCHQCTCIDAVAGCIPCLPQELSLPNIGCNPRILYKVTGGCCCEWVCDDESI 121
 OY 169 KDPMDODGLGKELGFPASEVELTRNNELLAVGKSLKRLPYGMEPRILYDLOGK 228
 DB 122 KDPMDODGLGKELGFPASEVELTRNNELLAVGKSLKRLPYGMEPRILYDLOGK 181
 OY 229 CIVOTTSWSQSKTCGTGISTRTVNDNECHRLVETRICELVPCGOPYSSLKRGKCSK 288
 DB 182 CIVOTTSWSQSKTCGTGISTRTVNDNECHRLVETRICELVPCGOPYSSLKRGKCSK 241
 OY 289 TKKSEPRFTYAGCLSKTKRPRKYGSCVGRCTPOLTRTYKMRFCED-----GE 341
 DB 242 TKKSEPRFTYAGCLSKTKRPRKYGSCVGRCTPOLTRTYKMRFCED-----GE 341
 OY 342 TFSKN---VMMIOSCKNVCNCPHANEAPFFYRLFNDFHFRD 381
 DB 296 IFGRHDDPVL-----KCNVNCPHANEAPFFYRLFNDFHFRD 334

RESULT 4
 ID 042607 PRELIMINARY; PRT: 343 AA.
 AC 042607;
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Connective tissue growth factor XCTGF.
 OS Xenopus laevis (African clawed frog).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
 CC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Ying Z., King M.L.;
 RL Submitted (Aug-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U43524; AAB67639.1; -
 DR EMBL; U43523; AAB67638.1; -
 DR InterPro: IPR006208; Cys_knot.
 DR InterPro: IPR006207; Cys_knot_C.
 DR InterPro: IPR000867; Insl_gro_fac.pr.
 DR InterPro: IPR000884; TSPL.
 DR InterPro: IPR001007; vWF-C.
 DR Pfam; PF00007; Cys_knot; 1.
 DR Pfam; PF00219; IGFBP; 1.
 DR Pfam; PF00090; tsp_1; 1.
 DR Pfam; PF00093; vwc; 1.
 DR SMART; SM00041; CT; 1.
 DR SMART; SM00121; IB; 1.
 DR SMART; SM00209; TSPL; 1.
 DR SMART; SM00214; VWC; 1.
 DR PROSITE; PS01185; CTCK_1; 1.
 DR PROSITE; PS01225; CTCK_2; 1.
 DR PROSITE; PS0092; TSPL; 1.
 DR PROSITE; PS01208; WFC; 1.
 SO SEQUENCE 343 AA; 37966 MW; 93F221C5DB565A81 CRC64;

Query Match 44.3%; Score 938; DB 13; Length 343;
 Best Local Similarity 47.2%; Pred. No. 5.8e-83;
 Matches 168; Conservative 51; Mismatches 95; Indels 42; Gaps 6;

OY 26 CPACHCPLEAPKCAPGVGLVDDGCGCCVCACQALNEDCSKTQPCDHTKLECNFGASST 85

DB 24 CNGEQCNKYPVCDPGRVWQDGGCCCKVSKQGLGELCTERDVCDPRKGLFCDRGSRVN 83
 OY 86 ALKICRAOSEGRPCENSRITQNGESQPNQCHQCTCIDAVAGCIPCLPQELSLPNIGC 145
 DB 84 RKIGVCTAR-EGAPCVFGFTYRSGESFOSCKYQCTCIDAVGCVPLCSMDIRLPSPDC 142
 OY 146 PNPRLYKVTGGCCCEWVCDDESIKDPMDODGLGKELGFPASEVELTRNNELLAVGGR 205
 DB 143 PPNRLYKVTGGCCCEWVCDDESIKDPMDODGLGKELGFPASEVELTRNNELLAVGGR 172
 OY 206 SLKRLPYGMEPRILYLPN---LOGQKCIYOTTSWSQSKTCGTGISTRTVNDNECHRLV 262
 DB 173 ---LPNARME---TYGDPDLIRANCLVQTTWNSAGSKTCGMGISTRTVNDNECHRLV 226
 OY 263 ETRICEVPCGOPYSSLKRGKCSKTKKSEPRFTYAGCLSKTKRPRKYGSCVGRCTPOL 322
 DB 227 QSRICMVRPCADIEENIKKRGKCIIRPKIRKPVKFEFGCTSVKTYRAKCGCTDGR 286
 OY 323 CTPOLTRTYKMRFCEDGFTSKNVMIOSCKNVCNCPHANE--AAPFFYRLFNDFI 376
 DB 287 CTPHRTATLPVEFKCPDGEVKNKMMFTKTCACHENCPGDDIFEAMITRKATGDM 342

RESULT 5
 ID 098T08 PRELIMINARY; PRT: 344 AA.
 AC 098T08;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Connective tissue growth factor precursor (Connective tissue growth factor/hypertrophic chondrocyte-specific protein 24).
 OS Gallus gallus (Chicken).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Gyi D.;
 RL Submitted (Jan-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Mukunda Y., Kubota S., Takigawa M.;
 RL Submitted (Dec-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ298335; CAC33438.1; -
 DR EMBL; AF463517; AAL68834.1; -
 DR InterPro: IPR006208; Cys_knot.
 DR InterPro: IPR006207; Cys_knot_C.
 DR InterPro: IPR000867; Insl_gro_fac.pr.
 DR InterPro: IPR000884; TSPL.
 DR InterPro: IPR001007; vWF-C.
 DR Pfam; PF00007; Cys_knot; 1.
 DR Pfam; PF00219; IGFBP; 1.
 DR Pfam; PF00090; tsp_1; 1.
 DR Pfam; PF00093; vwc; 1.
 DR SMART; SM00041; CT; 1.
 DR SMART; SM00121; IB; 1.
 DR SMART; SM00209; TSPL; 1.
 DR SMART; SM00214; VWC; 1.
 DR PROSITE; PS01185; CTCK_1; 1.
 DR PROSITE; PS01225; CTCK_2; 1.
 DR PROSITE; PS00222; IGF_BINDING; 1.
 DR PROSITE; PS0092; TSPL; 1.
 DR PROSITE; PS01208; WFC; 1.
 KW Signal.
 FT SIGNAL 1 21
 FT CHAIN 22 344
 SO SEQUENCE 344 AA; 37499 MW; 69E639A969BFLD00 CRC64;

Query Match 44.0%; Score 930; DB 13; Length 344;
 Best Local Similarity 45.3%; Pred. No. 3.5e-82;

	Matches	169;	Conservative	59;	Mismatches	105;	Indels	40;	Gaps	7;
OY		8	ALATVTLHLHTRALSTCPACHCPL-EAPKCAQVGILVRDCCGCCCKVCAYAKQLNEDCSK	66						
Dd		7	AVALLTALLIG-PEVAGGECGGCOCGSGPSPAGVSLVLVDGCSCCHVCAKQLGEELCTE	65						
OY		67	TQPCDHGTGLGNCRGASSTALKJGICORASEGRPCCYNSRITONGSRPDQNQHQTCLDG	126						
Dd		66	RDPGDHHHGGLCDRESPPANRRITGVCTAR-DAPCVFSQMWYRSGBSFFSSCKRYOCTCLDG	124						
OY		127	AVGCIFLCPDELSPINLGCNPNDRLVKYVGCCCEWVGCDEDSIKDMEMQDGLLGKELGFD	186						
Dd		125	AVGCVPLCMSVYRLPSDPCFPPRRKKLLGKCCEEWVCD-----	163						
OY		187	ASEVELTRNNELLIAVAGRSLSKLRL-PVEGMERILLYNLDLOKCIIVOTSMSSQSKTCGT	245						
Dd		164	-----AKBQTAAPALAAVRLIEDTVPDPPTM-----RANCLVQETEMSACSCTCGM	210						
OY		246	GISTRTVTNDNEBCRLVKEFRICEVARPCQPYYSSLKRRKSKTSKYKSEPRFYTAGCLS	305						
Dd		211	GISTRVTNDNAFCRIEKOSRCLCMVPCPADLEENIKKKKICIRPKISKPIFLSELGCTS	270						
OY		306	VKKTRPKYGSCVDRCTPOLFTFKRKRFRCDEGETSKNMVMIOSCKNVNCPHANEA	365						
Dd		271	VKTTRAKCGCTGORCTTPHRTATLPVEFKCPDGEIMKRMMFIKTACHINCPGDNDI	330						
OY		366	AFPFY-RLFNDI 376							
Dd		331	FESLYRRKMVGDM 343							
 RESULT 6 O9PT80 PRELIMINARY; PRT: 347 AA.										
ID	O9PT80	AC	O9PT80;							
DT	01-MAY-2000	(TREMBREL_13_Created)								
DT	01-MAY-2000	(TREMBREL_13_Last sequence update)								
DT	01-MAR-2003	(TREMBREL_23_Last annotation update)								
DE	Connective tissue growth factor,									
GN	CTGF.									
OS	Notophthalmus viridescens (Eastern newt) (Triturus viridescens).									
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;									
OC	Amphibia; Batrachia; Caudata; Salamandroidae; Salamandridae;									
OX	NCBI_TaxID=8316;									
RN	[1]									
RP	SEQUENCE FROM N.A.									
RA	TISSUE=Forelimb blastema;									
RL	Gates P.B.;									
RL	Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.									
RN	[2]									
RP	SEQUENCE FROM N.A.									
RC	TISSUE=Forelimb blastema;									
RX	MEDLINE=99033008; PubMed=9813273;									
RT	Cash D.E., Gates P.B., Imokawa Y., Brookes J.P.;									
RT	"Identification of new connective tissue growth factor as a target of									
RL	retinoid regulation in limb blastenal cells.";									
Gene	222:119-124(1998).									
DR	EMBL: AJ271167; CAB65965.1;									
DR	InterPro: IPR006208; Cys_knot									
DR	InterPro: IPR006207; Cys_knot.C									
DR	InterPro: IPR000867; Insl_gro_fac.pr									
DR	InterPro: IPR001007; WVF.C									
DR	Pfam: PF00007; Cys_knot.1.									
DR	Pfam: PF00219; IGFBP.1.									
DR	Pfam: PF00093; tsp_1; 1.									
DR	Pfam: PF00093; tsc_1; 1.									
DR	SMART: SMO0041; CT; 1.									
DR	SMART: SMO0121; IB; 1.									
DR	SMART: SMO0209; TSP1.1.									
DR	SMART: SMO0214; VWC; 1.									
DR	PROSITE: PS01185; CKCK_1; 1.									

DR	PROSITE:	PS01225;	CNCK_2; 1.
DR	PROSITE:	PS00222;	ICF_BINDING; 1.
DR	PROSITE:	PS50092;	TSP1; 1.
DR	PROSITE:	PS01208;	VWFC; 1.
SQ	SEQUENCE	347 AA;	38098 MW; 3B7E2399F27672C1 CRC64;
Query Match 43.1%; Score 912.5; DB 13; Length 347;			
Best Local Similarity 45.1%; Pred. No. 1,8e-80;			
Matches 167; Conservative 54; Mismatches 112; Indels 37; Gaps 6			
OY	9	LALVVTLLHLFTLSTLSTCPAACHCPLLEAPKCAPGVGLVBDGGCCGVCARAKOLNPEDSKTQ	68
Dd	12	LLIALLSVWSCA-QDSCGEGRCTNKRPPECAGATSLVNDGCGCCAYCAKKOLGELCTEKD	70
OY	69	PCHDTKGLECNFGASSTALKGICRAOSEGRPCENSRIVQNGESFPQNCOHOCTCIDAV	128
Dd	71	VODPHRGLFCDRGSVNNKKIGVCTAK-DGAPCVFGVMYRSGBSPGSSCKYCTCLDGCV	129
OY	129	GCIPLCPDEBLISPNIGCEPRLKYKTGGCCCEWVCDEDISIKPMEDQDLLEGELGFDS	188
Dd	130	GCVPPLCGVDVRLPSPDCPEPRRVKCOLGKCCEMVCDDPK-----	168
OY	189	EVELIRNNELLIAVGKGRSLKLRLPVGMERILYNPLQGKCIVOTSMGCSKTGTGIS	248
Dd	169	--EQTRVGPALAV-----YROETGYPD-----SSLMRANCLVOTTSMASCKTGMGIS	216
OY	249	TRVTDNNECRILKETRICEVAPCGOPYYSLSLKGGKCSKTKRSEPFVFYTAGLSVKR	308
Dd	217	TRVTDNNECMCREKERSRLCMVRPCADLEENIKKGRKCTRTPKISKPYAFELSGCSVAT	276
OY	309	YRPKYGSGVDRCCTPOLTRTVKRRFRCEDEGETFSKNVMTIQSCCKNYNCCHANPAEP	368
Dd	277	YRAKCGCYCTDGRCTCPHRTATLPLEVFCPCDDEVMAKKMKFIKTACAHYNCGDNDIFES	336
OY	369	FY--RLFNDI	376
Dd	337	MYYRKMYGDM	346
RESULT 7			
ID	097765	PRELIMINARY;	PRT; 349 AA.
AC	097765;		
DT	01-MAY-1999	(TREMBLrel. 10, created)	
DT	01-MAY-1999	(TREMBLrel. 10, last sequence update)	
DT	01-MAR-2003	(TREMBLrel. 23, last annotation update)	
DE	Connective tissue growth factor.		
GN	CTGF.		
OS	Sus scrofa (Pig).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.		
OX	NCHI_Taxid=9823;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=uterus.		
RA	Harding P.A., Briggsstock D.R.;		
RT	"Cloning and sequencing of a porcine connective tissue growth factor		
RT	(cDNA)." ;		
RL	Submitted (SEP-1996) to the EMBL/genbank/DDJB databases.		
DR	EMBL; U70060;	AAD00174.1; -	
DR	InterPro; IPR006208;	Cys_knot.	
DR	InterPro; IPR006207;	Cys_knot_C.	
DR	InterPro; IPR000867;	Insl_gro_fac_pr.	
DR	InterPro; IPR000884;	TSP1.	
DR	InterPro; IPR001007;	WMF_C.	
DR	Pfam; PF00007;	Cys_knot; 1.	
DR	Pfam; PF00219;	IGFBP; 1.	
DR	Pfam; PF00090;	tsp_1; 1.	
DR	Pfam; PF00093;	WC; 1.	
DR	SMART; SMART0041;	CT; 1.	
DR	SMART; SMART0121;	IB; 1.	
DR	SMART; SMART0209;	TSP1; 1.	
DR	SMART; SMART0214;	VMC; 1.	

DR PROSITE; PS50092; TSP1; 1.
DR PROSITE; PS01208; VWF; 1.
SQ SEQUENCE 367 AA; 40331 MW; 9F29CA94D69C0502 CRC64;

Query Match 36.2%; Score 767; DB 4; Length 367;
Best Local Similarity 38.0%; Pred. No. 2,8e-66;
Matches 148; Conservative 62; Mismatches 125; Indels 54; Gaps 9;

QY 2 SSRARALVYTLHLTRIALST-----CPACHCELEAPKCAPGVLRDGGCGCK 54
DB 18 STVALATSPAPPTMTFTFAPLEITREPECKMPCCECPSPPCPLGVSLITDGECK 77
QY 55 VCAKQNEQCKTPQPDHRTKGLCNF-GASSTALKGICRAGSEGRCEYNSRTYONGSEF 113
DB 78 MCAQOQGDCTEALCIDDPHRLGYCDYSGDRPRYAIGVC-AQVVGVCVLDGVRYYNGQSF 136
QY 114 QPNCQHCCTIDGAVGCIPLCPOELSLPNLGNPLRYKVTGGCCCEWYDCDESIKDPME 173
DB 137 QPNCCKNCTIDGAVGCIPLC-LRVPRPLMCPHPRVSTPGHCCQWVCEDDA-KRPRK 194
QY 174 DQDGLGKELGEPDA-SEVELTRNNELIAGKGRSLKRLPVGMEPRILYNPLQOKCIVQ 232
DB 195 TAPRDGT---AFDAVEVEAMHRN-----CIAY 219
QY 233 TTSWQCKTCGTGISTRTVNDNPECRLYKETRICEVRCQGPVYSSLKKGKCKSKTKS 292
DB 220 TTSWSPSTSGSGVSTRISVNAQCPQESRLCNLRPCDDVDIRHLIAGKCKLAVYQPE 279
QY 293 PEVPRVTYAGCLSVKKRYRKYCGSCVDGRCTPOLTRTVKMRFRCEDETFSKNMVMS 352
DB 280 EASMTTLAGCISTRTYRKYCGVCMNDNCIPYKSKTIDVSGQCDGLGFSNOVLMTNA 339
QY 353 CKCNVCPHANEAPFYRLFNDRHFRD 381
DB 340 CECNLSCRPNND-----IFADLESYPD 361

RESULT 10

099PPO PRELIMINARY; PRT; 367 AA.
AC 099PPO;
DT 01-JUN-2001 (TREMblrel. 17, Created)
DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)
DE ELM1.
GN ELM1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20487548; PubMed=11031104;
RA Sleeman M.A., Murlison J.G., Strachan L., Kumbale K.D., Glenn M.P.,
RA McGrath A., Bickerstaff P., Grierson A., Havukkala I., Tan P.,
RA Watson J.D.;
RT "Gene expression in rat dermal papilla cells: analysis of 2529 ESTs.";
RL Genomics 63;214-224(2000).
DR EMBL; AF228049; AAK00729.1; -
DR InterPro; IPR006208; Cys_knot.
DR InterPro; IPR006207; Cys_knot.C.
DR InterPro; IPR000867; Insl_gro_fac.Pr.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR001007; VWF_C.
DR Pfam; PF00007; Cys_knot; 1.
DR Pfam; PF00219; IGFBP; 1.
DR Pfam; PF00090; tsp_1; 1.
DR Pfam; PF00093; vwc; 1.
DR SMART; SM00041; CT; 1.
DR SMART; SM00121; IB; 1.
DR SMART; SM00209; TSP1; 1.
DR SMART; SM00214; VWC; 1.
DR PROSITE; PS01185; CTCK_1; 1.

DR PROSITE; PS01225; CTCK_2; 1.
DR PROSITE; PS50092; TSP1; 1.
DR PROSITE; PS01208; VWF; 1.
SQ SEQUENCE 367 AA; 40613 MW; 8AA4A34C69D3243D2 CRC64;

Query Match 35.8%; Score 758; DB 11; Length 367;
Best Local Similarity 38.5%; Pred. No. 2.1e-65;
Matches 149; Conservative 52; Mismatches 130; Indels 56; Gaps 8;

QY 5 IARALVYTLHLTRIALST-----CPACHCELEAPKCAPGVLRDGGCGCKKCA 57
DB 21 LATALSPTPTMTFTFAPLEITREPECKMPCCECPQAPRCPCLGVSLITDGECKICA 80
QY 58 KQNEQCKTPQPDHRTKGLCNF-GASSTALKGICRAGSEGRCEYNSRTYONGSEFQPN 116
DB 81 QOLGDCTEAAVCDPDRGLGYCDYSGDRPRYAIGVC-AQVVGVCVLDGVRYYNGSEFQPN 139
QY 117 COHCCTIDGAVGCIPLC--POELSLPNLGNPLRYKVTGGCCCEWYDCDESIKDPMD 174
DB 140 CRINCTIDGTGCTPLCLSPR---PLMCRQPRHVRVPGCCQWVCDDARR---PR 193
QY 175 QDGLGKELGEPDASEVELTRNNELIAGKGRSLKRLPVGMEPRILYNPLQOKCIVQT 234
DB 194 QTALDTRAFASGAVEREN-----CIAYNS 221
QY 235 SMSQCKTCGTGISTRTVNDNPECRLYKETRICEVRCQGPVYSSLKKGKCKSKTKSP 294
DB 222 PMSPCSTGGLGISTRTISVNAQCPQESRLCNLRPCDDVDIRHLIAGKCKLAVYQPE 281
QY 295 PVAFRTYAGCLSVKKRYRKYCGSCVDGRCTPOLTRTVKMRFRCEDETFSKNMVMS 354
DB 282 ATNFTLAGCISTRTYRKYCGVCMNDNCIPYKSKTIDVSGQCDGLGFSNOVLMTNCF 341
QY 355 CYNVCPHANEAPFYRLFNDRHFRD 381
DB 342 CNLSCRPNND-----IFADLESYPD 361

RESULT 11

054775 PRELIMINARY; PRT; 367 AA.
AC 054775;
DT 01-JUN-1998 (TREMblrel. 06, Created)
DT 01-JUN-1998 (TREMblrel. 06, Last sequence update)
DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)
DE ELM1.
GN ELM1 OR WISPL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=HEN;
RX MEDLINE=98119879; PubMed=9449709;
RA Hashimoto Y., Shindo-okada N., Tani M., Nagamachi Y., Takeuchi K.,
RA Shiroishi T., Toma H., Yokoyama J.;
RT Expression of the ELM1 gene, a novel gene of the CCN (connective
RT tissue growth factor, Cy61/Ce10, and neuroblastoma overexpressed
RT gene) family, suppresses in vivo tumor growth and metastasis of K-1735
RL murine melanoma cells.";
RL J. Exp. Med. 187:289-296(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Mammary gland;
RX MEDLINE=99061933; PubMed=9843955;
RA Penica D., Swanson T.A., Welsh J.W., Roy M.A., Lawrence D.A., Lee J.,
RA Brush J., Taneyhill L.A., Quirk B., Lew M., Watanabe C., Cohen R.L.,
RA Melham M.F., Finley G.G., Quirk P., Goddard A.D., Hillan K.J.,
RA Gurney A.L., Botstein D., Levine A.J.;
RT "Wisp genes are members of the connective tissue growth factor family
RT that are up-regulated in wt-1-transformed cells and aberrantly
RT expressed in human colon tumors.";

Proc. Natl. Acad. Sci. U.S.A. 95:14717-14722(1998).
 DR EMBL; AB004873; BAA24949.1; -
 DR EMBL; AF100777; AAC96319.1; -
 DR MGD; MGI:1197008; Wtsp1.
 DR InterPro; IPR006208; Cys_knot.
 DR InterPro; IPR006207; Cys_knot_C.
 DR InterPro; IPR000867; Insl_gro_fac_pr.
 DR InterPro; IPR000884; TSP1.
 DR InterPro; IPR001007; WVE_C.
 DR Pfam; PF00007; Cys_knot; 1.
 DR Pfam; PF00219; IGFBP; 1.
 DR Pfam; PF00090; tsp_1; 1.
 DR Pfam; PF00093; wvc; 1.
 DR SMART; SM00041; CT; 1.
 DR SMART; SM00209; TSP1; 1.
 DR SMART; SM00214; IB; 1.
 DR SMART; SM00214; VWC; 1.
 DR PROSITE; PS01185; CTCK_1; 1.
 DR PROSITE; PS01225; CTCK_2; 1.
 DR PROSITE; PS50092; TSP1; 1.
 DR PROSITE; PS01208; WVEC; 1.
 DR SEQUENCE 367 AA; 40702 MW; 387C0569EFAE5E96 CRC64;
 Query Match 35.8%; Score 758; DB 11; Length 367;
 Best Local Similarity 38.5%; Pred. No. 2,1e-65;
 Matches 149; Conservative 52; Mismatches 130; Indels 56; Gaps 8;
 QY 5 IARALAVTLHLTRALST-----CPACHCLEAPKCAPGVGLVRDGGCCGVCA 57
 Db 21 IATALSTPTPTMTTPAPLEETTRPFCCKWPCPCPSPPCPLGVSILTDGCECKICA 80
 QY 58 KOLNEDSKTQPCDHTGLECNF--GASTALKGICRQSESRPEVNSRIYQNESEOPN 116
 Db 81 QQLDNCETEAICDPHGLCYDSGDRPRAIGVC-AQVVGCVLLDGVRTNGESFOPN 139
 QY 117 COHCCITIDAVNGCIPLC--POELSLPNLGCNPNRLVKGCCCEWVCDEDSIKDPMED 174
 Db 140 CRVCTCIDGVTGCTPLCLSPR---PRMCRQPRHVRVPGCCCEWVCDDAR----PR 193
 QY 175 QDGLGKELGDFDASEVELTRNNELIYAVGKSKRLRLEPFVGNPRILYNPLQGGKCIQVOT 234
 Db 194 QTALLDTRAPFAASGAVQREYN-----CIAYTS 221
 QY 235 SWSCGCTGCTGISTRYTNNPECRLYKEPRICEVRPGCGPVYSLKGRKCKSTKSPE 294
 Db 222 PWSGCTGCTGISTRYTNNPECRLYKEPRICEVRPGCGPVYSLKGRKCKSTKSPE 281
 QY 295 PVRETYAGCLSVKRYKYGSCVDGRCTPOLTRTYKMRPCEDEGETFSKNVMIOGSK 354
 Db 282 ATNFTLAGCVSTRYRKYRKYGVCTDNRCICPYKSKTISVDQCPGEGFGFSNOVIMINACF 341
 QY 355 CNVNCPRANEAAPFYRLNDIHKFRD 381
 Db 342 CNLSCRNPND-----IFADLESYPD 361
 RESULT 12
 095958 PRELIMINARY; PRT; 331 AA.
 AC 095958;
 DT 01-MAY-1999 (Tremblrel. 10, Created)
 DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
 DE D14217.3 (Connective tissue growth factor (NOV, GIG) like protein)
 DE (Fragment).
 GN D14217.3 OR LIBC.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Tubby B.;

Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 RN 12
 RP SEQUENCE FROM N.A.
 RC TISSUE=Mammary gland;
 RA van Golen K.L., Davies S., Wu Z.F., Wang Y.F., Bucana C.D., Root H.,
 RA Chandrasekharappa S., Strawderman M., Ehlier S.P., Merajver S.D.,
 RT "A novel putative IGF-binding, tumor suppressor protein, LIBC, and
 RT Rhc GTPase, are determinants of the inflammatory breast cancer
 RT phenotype";
 RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; Z99289; CAB15556.1; -
 DR EMBL; AF143679; AAD31517.1; -
 DR InterPro; IPR006208; Cys_knot.
 DR InterPro; IPR006207; Cys_knot_C.
 DR InterPro; IPR000867; Insl_gro_fac_pr.
 DR InterPro; IPR000884; TSP1.
 DR Pfam; PF00007; Cys_knot; 1.
 DR Pfam; PF00219; IGFBP; 1.
 DR Pfam; PF00090; tsp_1; 1.
 DR SMART; SM00041; CT; 1.
 DR SMART; SM00209; TSP1; 1.
 DR SMART; SM00214; IB; 1.
 DR SMART; SM00214; VWC; 1.
 DR PROSITE; PS01225; CTCK_2; 1.
 DR PROSITE; PS00222; IGF_BINDING; 1.
 DR PROSITE; PS50092; TSP1; 1.
 DR NON TER 331
 DR SEQUENCE 331 AA; 36909 MW; D109C2FDCAD1DF549 CRC64;
 Query Match 29.6%; Score 626.5; DB 4; Length 331;
 Best Local Similarity 36.2%; Pred. No. 1.2e-52;
 Matches 122; Conservative 50; Mismatches 118; Indels 47; Gaps 8;
 QY 26 CPACHCLEAPKCAPGVGLVRDGGCCGVCAKQINDECSCTQCDHTKGLGECNFGASST 85
 Db 26 CHMCKCPCKPCPCPSPPCPLGVSILTDGCECKICAQPGELCNADLCPKGYCYSDRP 85
 QY 86 ALK-GICRAOSEGRPEVNSRIYQNESEOPNCOHCTCIDAVNGCIPLCQELSLPNLG 144
 Db 86 RYETGVCAKSVG--CERNVYHNGVFOQNPFLPCLCYSGAIGCTPLF----- 133
 QY 145 CPNRLVKGCCCEWVCDEDSIKDPMEDDGLGKELGDFDASEVELTRNNELIYAVGK 204
 Db 134 -----IPKLGSHC-----SGAKGGRKS-DQNSCSL-----EPDQGLS 166
 QY 205 RSLKRLVFGMEPRILYNPLQGGKCIQVOTSWSCGCTGCTGISTRYTNNPECRLYKET 264
 Db 167 TSYKTPAYRNLPPI-----WKKKLVQATKTPCSRTGCGISNRYTNNENSCMRKEX 221
 QY 265 RICEVRPGCGPVYSLK--KSKCKSTKSPKPEVRYAGCLSVKRYKYGSCVDGRG 322
 Db 222 RLCTIQPDSDNLTITIKIPKCKTQPTFOLSKAEVFGSGSSYQSTKPTFGICLDKRC 281
 QY 323 CTPOLTRTYKMRPCEDEGETFSKNVMIOGSKNYNC 359
 Db 282 CIPNKSMTITIQFDCPNBGSFKMKMLMTTSCVQRC 318
 RESULT 13
 095389 PRELIMINARY; PRT; 354 AA.
 AC 095389;
 DT 01-MAY-1999 (Tremblrel. 10, Created)
 DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
 DE Connective tissue growth factor related protein WISP-3.
 DE WISP3.
 GN WISP3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA TISSUE=Bone marrow, and fetal kidney;

RX MEDLINE-99061933; PubMed-9843955;
 RA Pennica D., Swanson T.A., Welsh J.W., Roy M.A., Lawrence D.A., Lee J.,
 RA Brush J., Taneyhill L.A., Deuel B., Lew M., Watanabe C., Cohen R.L.,
 RA Melham M.F., Finley G.G., Quirke P., Goddard A.D., Hillan K.J.,
 RA Garney A.L., Botstein D., Levine A.J.;
 RT "WISP genes are members of the connective tissue growth factor family
 that are up-regulated in wnt-1-transformed cells and aberrantly
 RT expressed in human colon tumors."
 RL Proc. Natl. Acad. Sci. U.S.A. 95:14717-14722(1998).
 DR EMBL: AF100781; AAC96323.1; -
 DR Genew: HGNC:12771; WISP3.
 DR InterPro: IPR006208; Cys_knot.
 DR InterPro: IPR006207; Cys_knot_C.
 DR InterPro: IPR000867; Insi_gro_fac_pr.
 DR InterPro: IPR000884; TSP1.
 DR Pfam: PF00007; Cys_knot; 1.
 DR Pfam: PF00219; IGFBP; 1.
 DR Pfam: PF00090; tsp_1; 1.
 DR SMART: SM00041; CT; 1.
 DR SMART: SM00121; IB; 1.
 DR SMART: SM00209; TSP1; 1.
 DR PROSITE: PS01225; CTCK; 2; 1.
 DR PROSITE: PS00222; IGF_BINDING; 1.
 DR PROSITE: PS50992; TSP1; 1.
 SQ SEQUENCE 354 AA; 38292 MW; 67F48DD5C2F5EE3 CRC64;

Query Match 29.0%; Score 613; DB 4; Length 354;
 Best Local Similarity 35.9%; Pred. No. 2,6e-51;
 Matches 121; Conservative 49; Mismatches 121; Indels 46; Gaps 8;

QY 26 CPACACPLEAPKAGVGLVNDGCGCCVCAKQLNEDSKTOPCHTKLEGNEFASST 85
 DB 48 CHWPCKCPDQKPRCPGVSLVRDGCCKKCAKQPEICNEALCDPHKGLCYDYSDRP 107
 QY 86 ALK-GICRAQSEGRPEYNSRITQNGESFPQNCQHCTCIDGAVGICPLCPQELSLPNI 144
 DB 108 RYETGVC-AYLVAAGEFENYVHNQVOPNPLFSLCYSGAIGCTPLF----- 156
 QY 145 CPNPLRVLYTGCCCEEWVDEDSIKDPMEDDGLLKEGLFASVEVELRNELLAVGKG 204
 DB 157 -----IPKLAGSHC-----SGAKGGRKS-DQSNCSL-----EPLIQQLS 189
 QY 205 RSLKRLPVFGMERILYNPLQGGKCIYQTTSMQCKTGCTGISTRYTNDNPRCLVKET 264
 DB 190 TSKTKNPAYRNPLI-----WKKKCLVQATKWPCTGCMGINSYTNENSCENRKEK 244
 QY 265 RICEVPPCGOPYSSLSK-KGKKCKSKKSPPEVPTVYAGLSVKKYRPYCGSCVDGRG 322
 DB 245 RLCTIOPCDSNIIKTKIKRGTCTPTPOLSKAEKVFVSGCSSTQSKTFFGICLDRK 304
 QY 323 CTPQLRTVYKMRFCEDGETFSKNVMQSCKQNYNC 359
 DB 305 CIPNRSKMITIOPDCPNESGFKWKMIMITSCVQQRNC 341

RESULT 14

ID 08CIC8 PRELIMINARY; PRT; 251 AA.
 AC 08CIC8;
 DT 01-MAR-2003 (Tremblrel. 23, Created)
 DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
 DE Similar to WNT1 inducible signaling pathway protein 2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Mammary gland;
 RA Strausberg R.;
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC032877; AAH32877.1; -

SQ SEQUENCE 251 AA; 27106 MW; 57520309A9069524 CRC64;

Query Match 24.3%; Score 514.5; DB 11; Length 251;
 Best Local Similarity 36.7%; Pred. No. 6,8e-42;
 Matches 99; Conservative 36; Mismatches 86; Indels 49; Gaps 5;

QY 15 LHLTRLAIS-----TCPACACPLEAPKAPGVLVNDGCGCCVCAKQLNED 63
 DB 6 LHL--LAISFLCILSMVYAOQLCPAPCAPWFPQCPGPVPLVLDGCGCCVCAKRLGES 63
 QY 64 CSKTOPCDHTKLEGNEFASSTALKGICRAQSEGRCEYNSRITQNGESFPQNCQHCTC 123
 DB 64 CDHLVACNPQGLVCPGAGPSGRGVCLPEEDGSCVNGRRYLDGETFKPCRYLCRC 123
 QY 124 IDGAVGICPLCPQELSLPNIIGCPNPLVAVTGCCCEEWVDEDSIKDPMEDDGLLKE 183
 DB 124 DDGFLCTPLCSADVRLPMSDCRRRIQVPGKCGCEWCDQ-AVMQPAIOLSSAQGH 182
 QY 184 GPDASEVELTRNNELLAVGKSLKRLPVGMEPRILYNPLQGO-KCIYQTTSMQCKST 242
 DB 183 S-----ALVTPASADGCPMWMSTAMGCPSTT 208
 QY 243 CGGTSTRYTNQDPECRIVKETRICVPRPC 272
 DB 209 CGLGIAFRVSNQNRFCQLEIQRRLCLSRPC 238

RESULT 15

ID 09HCS3 PRELIMINARY; PRT; 280 AA.

AC 09HCS3;
 DT 01-MAR-2001 (Tremblrel. 16, Created)
 DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
 DT 01-MAR-2001 (Tremblrel. 23, Last annotation update)
 DE WISP-1 variant.
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Tanaka S., Sugimachi K.;
 RT "Human WISP-1 variant."
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB034725; BAB17849.1; -
 DR InterPro: IPR006208; Cys_knot.
 DR InterPro: IPR006207; Cys_knot_C.
 DR InterPro: IPR000867; Insi_gro_fac_pr.
 DR InterPro: IPR000884; TSP1.
 DR Pfam: PF00007; Cys_knot; 1.
 DR Pfam: PF00219; IGFBP; 1.
 DR Pfam: PF00090; tsp_1; 1.
 DR SMART: SM00041; CT; 1.
 DR SMART: SM00121; IB; 1.
 DR SMART: SM00209; TSP1; 1.
 DR PROSITE: PS01185; CTCK; 1; 1.
 DR PROSITE: PS01225; CTCK; 2; 1.
 DR PROSITE: PS50992; TSP1; 1.
 SQ SEQUENCE 280 AA; 30697 MW; 26B254DA060738E CRC64;

Query Match 23.4%; Score 495.5; DB 4; Length 280;
 Best Local Similarity 27.9%; Pred. No. 5,4e-40;
 Matches 108; Conservative 49; Mismatches 93; Indels 137; Gaps 8;

QY 2 SSRIARALALVYTLHLTRLAIST-----CPACACPLEAPKAPGVLVNDGCGCC 54
 DB 18 SYVALTALSPAPRTMDFTFPAPLEDTSRQFCKWPCCEPPSPRCLGVSLITDCECK 77
 QY 55 VCAKQLNEDSKTOPCDHTKLEGNEFASSTALKGICRAQSEGRCEYNSRITQNGESFP 114
 DB 78 MCAQQLGDNCTEALICDPHRLGYCDY-----SGDRP-----RYAIGV--- 114
 QY 115 PNCQHCTCIDGAVGICPLCPQELSLPNIIGCPNPLVAVTGCCCEEWVDEDSIKDPMED 174

```
Db      115  --CAH-----AVG-----120
QY      175  QDGLGKELGFDASEVELTRNNELIavgKGRSLKRLPVGMEPRILYNPLQGQCIvQTT 234
Db      121  -----EVEAMHRN-----CIAVTS 134
QY      235  SWSGCKTGIGTSTRYTNDNPEGRVKEPRICEVRPGQPVYSSLKKGKCKSKTKKSPE 294
Db      135  PWSPCSTSCGLGVSTRISNVAOCWPEQESRLCMLRPCDDVDIHTLIKAGKCLAVYOPEA 194
QY      295  PVRETYAGCLSVKKYRPRKYGSCYDGRCTPOLTRTYKMRPRCEDEGTFESKNVMMIOSCK 354
Db      195  SMNFTLAGCISTRSYOPKYGCVCMNRCCLPYKSKTIDVSRQCPDGLGFSRQVLMINACF 254
QY      355  CNYNCPHANEAPFPYRLFENDIHKFRD 381
Db      255  CNLSCRNPND-----IFADLESYPD 274
```

Search completed: August 5, 2003, 14:06:51
Job time : 34.0868 secs

THIS PAGE BLANK (USPTO)